THE UNIVERSITY OF ALBERTA

Second Language Acquisition: A Developmental Study of Comprehension, Production, and Semantic Memory Organization

by

Mary Susan Crawford

(C)

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1986

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-32475-9

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Mary Susan Crawford

TITLE OF THESIS Second Language Acquisition: A Developmental Study of Comprehension, Production, and Semantic Memory Organization

DEGREE FOR WHICH THESIS WAS PRESENTED DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED FALL, 1986

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) Mary Sue Crawford

PERMANENT ADDRESS:

262 Rose St. #307

BARRIE, ONTARIO

L4M 2V4

DATED SEPTEMBER 23 19[illegible]

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled Second Language Acquisition: A Developmental Study of Comprehension, Production, and Semantic Memory Organization submitted by Mary Susan Crawford in partial fulfilment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

.....

Supervisor

.....

.....

.....

.....

External Examiner

Date SEPTEMBER 23, 1986

Abstract

A cross-sectional method was used to study the development of comprehension, production, and organization of semantic memory among second language learners. Canadian anglophone students attending either a French immersion or extended French program were studied at grades 6 and 9. The choice of these four groups made it possible to examine effects associated with duration of exposure and age (Grade) and effects associated with intensity of exposure (Program). Production was assessed with two tasks, picture and word naming, and comprehension was assessed with one task, picture-word matching. The same single-word stimulus items were used in all three tasks. A bilingual version of a category matching task was used to examine organization of semantic memory. Bilingual balance was achieved earlier on the comprehension task, than on either of the two production tasks. For students in both programs, French response latencies decreased across grade. Of interest was whether there was a corresponding decrease in English response latencies across grade. The results differed for students in each program. Among extended French pupils, there was a significant decrease in word-naming and picture-naming matching latencies. In contrast, among immersion students, there was a marginally significant decrease in picture-word matching, but a marginally significant increase in word-naming. In the category matching task, it was found that, irrespective of whether comprehension latencies were equivalent in the two languages, category knowledge for both languages was represented in a common memory store. The results

are interpreted as support for a one-store model of memory, and the pattern of first and second language comprehension and production latencies is accounted for by increased automaticity and decreased interference between the two languages.

Acknowledgements

I would like to thank Dr. Jeffrey Bisanz for his many hours of help, his invaluable comments and criticisms at each stage of thesis, and also for his continued support throughout.

I would like to express my appreciation to Dr. Steve Carey for sharing his expertise in second language acquisition with me. From the initial development of the ideas for the thesis to the final draft, his comments have been extremely valuable.

Dr. Ed Cornell's comments about developmental and methodological issues were very helpful. I would like to take this opportunity to thank him for his constructive criticism throughout each stage of the thesis.

Dr. Peter Dixon agreed to be a committee member in the spring of 1986, following Dr. Spearman's retirement. I would like to thank him for his involvement at this late stage and for the many helpful comments he offered, particularly in the area of statistics.

A special note of thanks to Hau Chow for her support throughout, for all her help collecting data, and for the numerous times she ran off the hard copy of my thesis and sent it to me while I was doing a clinical internship in Ontario. I want to thank Bob Mitchnick, whose encouragement and understanding kept me going through times of frustration. Also, I want to thank him for the many hours he spent working on the figures.

Table of Contents

Chapter — Page

I. INTRODUCTION 1

A. Development of Comprehension and Production in Two Languages 7

Encoding and Decoding 7

Bilingual Balance 11

Effects of Exposure 14

Verbal Intelligence 20

Differential Rates of Skill Development 22

Facilitation Effects 24

B. Organization of Bilingual Semantic Memory 26

Description of the Category Matching Task 33

C. Summary of the Issues 39

II. METHOD 41

A. Subjects 41

B. Materials and Procedures 42

III. RESULTS AND DISCUSSION 49

A. Verbal Intelligence Scores 49

B. Accuracy Data 54

C. Picture-Naming Latencies 54

D. Word-Naming Latencies 67

E. Picture-Word Differential Latencies 79

F. Picture-Word Matching Latencies 79

G. Category Matching Latencies 88

IV. GENERAL DISCUSSION 100

A. Effects of Exposure 101

B. Verbal Intelligence 106

C. Differential Rates of Skill Development 108

D. Facilitation Effects 111

E. Organization of Bilingual Semantic Memory 112

V. CONCLUSIONS 115

VI. REFERENCES 116

VII. APPENDIX A: COMPREHENSION AND PRODUCTION STIMULI NORMS 129

VIII. APPENDIX B: PICTORIAL STIMULI 132

IX. APPENDIX C: WORD NORMS FOR CATEGORY MATCHING 136

X. APPENDIX D: WORD LISTS FOR CATEGORY MATCHING 138

XI. APPENDIX E: ACCURACY DATA 140

List of Tables

Table	Description	Page
1	Predictions for Latency Data from the Category Matching Task	37
2	Mean Verbal IQ Scores for French Immersion and Extended French Pupils	50
3	Correlations Between Verbal IQ Scores and French, English and French-English Differential Latencies for Picture-Naming, Word-Naming and Picture-Word Matching Tasks	53
4	Mean Latencies and Accuracy in Percentages for the Picture-Naming (PN), Word-Naming (WN), and Picture-Word Matching (PWM) Task	55
5	Percentage of Subjects with Incomplete Latency Data or Negative Intercepts for Picture-Naming Task	57
6	French, English and French-English Differential Latencies for the Picture-Naming, Word-Naming, and Picture-Word Matching Tasks	62
7	Mean English Latency Differences and Confidence Intervals for Picture Naming, Word Naming, and Picture-Word Matching	66
8	Percentage of Subjects with Incomplete Data or Negative Intercepts for the Word-Naming Task	69
9	Picture-Naming, Word-Naming and Differential Picture-Word Latencies	80
10	Median Latencies and Accuracy Scores for the Category Matching Task	89
11	Correlations Between Verbal IQ and Category Matching Latencies	90
12	Observed Values for Prediction Equations for Models of Bilingual Semantic Memory	94
13	Median Latencies and Accuracy Scores for Grade 9 Immersion Students on the First and Second Half of the Category Matching Task	98

List of Figures

Figure	Description	Page
1	A schematic diagram of a model for picture and word processing	9
2	One-Store and Two-Store models of bilingual semantic memory	36
3	Mean response latencies for naming pictures as a function of Grade and Language	59
4	Mean response latencies for naming pictures as a function of Program and Language	61
5	Mean response latencies for naming pictures as a function of Grade, Program, and Language	63
6	Mean response latencies for reading words aloud as a function of Grade and Language	71
7	Mean response latencies for reading words aloud as a function of Program and Language	72
8	Mean response latencies for reading words aloud as a function of Grade and Program	73
9	Mean response latencies for reading wrods aloud as a function of Grade, Language and Testing Sequence	75
10	Mean response latencies for reading words aloud as a function of Program, Sex, and Task Order	77
11	Mean response latencies for reading words aloud as a function of Program and Practice across trials	82
12	Mean response latencies for matching picture-word pairs as a function of Testing Sequence and Language	84
13	Mean response latencies for matching picture-word pairs as a function of Program and Language	85
14	Mean response latencies for all groups at each language condition for matching trials on the Category Matching Task	89
15	Mean response latencies for French Immersion Grade 9 pupils on the matched trials of the Category Matching Task as a function of Language Condition and Trial	96

I. INTRODUCTION

In the past decade, there has been a proliferation in the number and variety of French programs available for English-speaking students in Canada (e.g., Swain & Lapkin, 1981). Exposure to the second language in these programs varies from intensive and early, to minimal and later. An example of the former type is the early French immersion program. Students are instructed only in French from kindergarten until grade 2, at which time English-instruction is introduced for one subject and gradually increased to include about 40% of the curriculum by grade 9. In contrast, the extended French program exemplifies the approach with minimal and later exposure. Students are instructed only in English from kindergarten to grade 4, at which time they receive three hours of French instruction per week. The amount of French instruction is increased gradually between grades 4 and 9 to include one additional subject taught in French (e.g., social studies).

Early research on French language programs in Canada was focused on applied issues, in particular comparison of achievement levels in French and English language skills (e.g., Lambert & Tucker, 1972; Bruck, Lambert & Tucker, 1974; Swain & Barik, 1976). Recently, the focus has been extended to include examination of the processes involved in acquiring a second language, such as comprehension and production. Comprehension refers to the ability to understand either written or spoken language, whereas production refers to the ability to produce oral or written language (see, Clark & Clark, 1977; deVilliers &

deVilliers, 1978). Another important issue, but one that typically is addressed separately from comprehension and production, is the question of how semantic knowledge from two languages is organized in memory (e.g., Caramazza & Brones, 1980; Kolers & Paradis, 1980; Snodgrass, 1984).

Three central questions are addressed in the present study. First, how is response speed on simple comprehension and production tasks, in the first and second language, affected by differences in exposure to the second language? Second, how is the organization of semantic knowledge in memory affected by exposure to the second language? Third, what is the relationship between comprehension speed in each language and organization in bilingual memory? The population of interest is Canadian anglophone students attending either a French immersion or extended French program.

The majority of prior research on the development of second language comprehension and production skills was conducted in Europe (e.g., Dornic, 1979; Magiste, 1979, 1982, 1984), where students have highly diversified linguistic backgrounds. As a group of second language learners, Canadian anglophone students differ markedly from their European counterparts. In many parts of Canada, students are relatively homogeneous with respect to linguistic background in that the majority speak English as their first language. Within immersion or extended French programs, most students are at the same age when second language instruction begins and the amount of exposure to the second language is relatively similar because there is limited

additional exposure to French outside the school environment. There are exceptions, such as in the province of Quebec, where French is the dominant language (the language most frequently used). Given the difference between the European and Canadian student population, it cannot be assumed that developmental trends reported for European students will be replicated. The trend of particular interest was that reported by Magiste (1979, 1982). She found that as the number of years of exposure to the second language increased, response latencies on simple comprehension and production tasks decreased in the second language, but increased in the first language.

The present sample was comprised of four groups of Canadian anglophone students: grade 6 and 9 immersion and extended French pupils. This cross-sectional design made it possible to examine two different aspects of exposure to a second language. First, effects associated with intensity of exposure can be assessed by comparing subjects in immersion (intense exposure) and extended French (minimal exposure) programs. Of particular interest, is whether students with more exposure to the second language take longer when working in their first language than students with minimal exposure. Second, effects of duration of exposure (i.e., number of years) can be assessed by comparing response latencies across grade. In the present study, duration of exposure and age are confounded. Thus the question arises as to whether the results reflect an increase in first language response latencies, associated with greater exposure to the second language (e.g., Magiste, 1979, 1982), or a decrement in response speed typically

associated with increasing age (e.g., Goodman, Haith, Guttentag, & Rao, 1984). A change in response latencies across grades that differs in size or direction for immersion and extended French pupils would suggest that the results do not merely reflect age effects because age effects should be the same for both groups. What such a pattern of results would suggest is that there is an effect of duration that varies as a function of intensity of exposure. That is, an increase in English response latencies across grades might be found only among students receiving intense exposure to French.

In previous studies, a control group of monolinguals often has been included. In the present study, the choice was made to include a group with minimal exposure to French, grade 6 extended French pupils, rather than monolingual group, for several reasons. First, the primary focus of the present study was on the development of second language skills and the critical comparisons were between level of skill in each language. To make these comparisons, subjects needed to be able to work in both languages. Second, inclusion of a group of subjects with mimimal exposure to French was thought to more accurately reflect the Canadian linguistic context and thereby ensure that the research questions addressed would have more direct applicability.

Verbal intelligence, as measured on standardized intelligence tests, was another variable included in the present design. Although it was not possible to equate completely for Verbal IQ when selecting subjects, statistical control was

possible. Furthermore, collection of Verbal IQ data facilitated examination of the relationship between Verbal IQ and comprehension and production skills in a second language.

The choice of comprehension and production tasks is another aspect of the present study that warrants mention. In previous studies, it has been difficult to draw conclusions about the relative rate of development of second language comprehension and production skills for two reasons. First, the stimuli were seldom equated for important characteristics such as frequency. Second, the level of task difficulty has differed for the comprehension and production tasks (e.g., Dornic, 1979; Magiste, 1979). In the present study, production was assessed with two tasks, picture and word naming, and comprehension was assessed with one task, picture-word matching. The same, single-word stimulus items were used in all three tasks. Thus it was possible to compare the relative rate of development of these specific comprehension and production skills without possible confoundings due to stimulus materials or task difficulty.

The focus in the present study was on the effects of duration and intensity exposure on the organization of bilingual semantic memory, as well as comprehension and production. In the research on bilingual semantic memory, the major focus has been on whether bilingual processing is more appropriately conceived of as involving one store for both languages, or two separate stores (see Glucksberg, 1984, for a review of these two models). Although numerous studies have been published in the last 20 years, the controversy has not been resolved (e.g., Caramazza &

Brones, 1980; Dryer, 1971; Durga, 1978; Goggin & Wickens, 1971; Kintsch, 1970; Kolers, 1966a; Kolers & Gonzalez, 1980). Recently, it has been hypothesized that bilingual semantic memory may undergo structural change as a function of years of exposure to the second language (Gekoski, 1980; Potter, So, Eckhart, & Feldman, 1984). It has been proposed that when subjects have had minimal exposure to the second language and are much more fluent in their first language, such as the grade 6 extended French pupils, they may utilize two semantic stores and translate between them. However, among subjects who are equally fluent in both languages, semantic knowledge may be stored in one memory. The purpose of the present study was to investigate (a) whether subjects stored semantic knowledge from their two languages in one or two stores, and (b) whether this aspect of semantic memory organization changes as a function of intensity or duration of exposure to French. To this end, the same four groups of subjects were assessed on a bilingual category matching task that involves semantic processing.

In summary, the focus of the present study is on the effects of exposure to a second language on comprehension, production, and organization of semantic memory. To assess effects of exposure, the sample included four groups of subjects differing with respect to duration and intensity of exposure to French. To investigate comprehension, production, and organization of memory, all subjects were tested in French and English on a picture-naming, word-naming, picture-word matching, and category matching task.

In the following section, research is described that pertains to the measurement and development of comprehension and production skills among second language learners. Research on the organization of bilingual semantic memory is then reviewed.

A. Development of Comprehension and Production in Two Languages

Encoding and Decoding

As defined by psycholinguists studying first language acquisition (e.g., Clark & Clark, 1977; deVilliers & deVilliers, 1978), production refers to the ability to generate meaningful verbal output. Comprehension refers to the ability to see or hear a word and identify nonverbally the meaning associated with the visual or auditory stimulus.

In information-processing models, the distinction between comprehension and production skills has been interpreted by some authors in terms of encoding and decoding operations, respectively (Dornic, 1979; Magiste, 1979, 1982). Encoding generally refers to the process by which external stimuli are perceived and translated into internal codes. Consider, for example, Snodgrass' (1984) model of picture and word processing, as illustrated in Figure 1. Snodgrass postulates that at Level I, subjects process such physical characteristics as typeface and size (printed words), orientation and amount of detail (pictures), or voice and intonation (spoken words). At Level II, subjects translate the stimuli into prototypical acoustic or visual images. Depending on the level of processing required by

the task, stimuli may be encoded at Level III, the propositional/semantic store. To assess whether a stimulus has been encoded, a response is required. A typical requirement of a task measuring comprehension/encoding skills is that the response be nonverbal.

Insert Figure 1 about here

Decoding generally refers to the process by which internal codes are translated into external codes. The unique requirement of production tasks is that the stimuli must be translated into verbal codes. Using Snodgrass's (1984) model, it appears that the stimulus can be decoded at either the propositional/semantic store (Level III), or the prototypical image store (Level II), depending on task demands. With respect to the demands in picture- and word-naming tasks, some specific predictions have been developed. Words are named some 200 ms to 300 ms faster than pictures (e.g., Potter & Faulconer, 1977; Potter et al., 1984). Within Snodgrass's (1984) model, "this effect is accounted for by assuming that pronouncing a word requires access only to Level II, with no necessity to access the deeper semantic level. For pictures, in contrast, access to the semantic level is required before accessing the name code for the picture, thereby requiring more processing and more time" (p. 18). It should be noted that in both production tasks, the stimulus must be encoded at an appropriate level before it can be decoded and articulated.

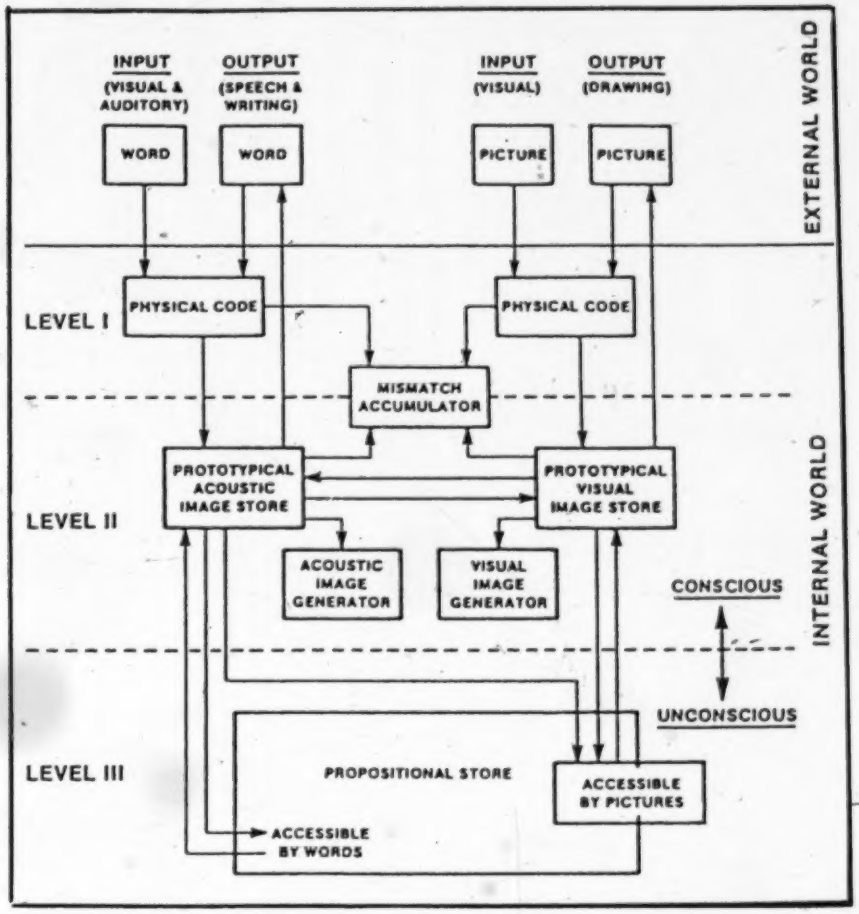


Figure 1. A schematic diagram of a model for picture and word processing. (Note: From "Concepts and their surface representations" by J. G. Snodgrass, 1984, Journal of Verbal Learning and Verbal Behavior, 23, p.16.)

Based upon this aforementioned distinction between processing pictures and words, it can be assumed that, in the picture-word matching task, pictures are encoded initially at Level III and words at Level II. Assuming that codes for the two stimuli must be at the same level if a subject is to respond accurately, the word must be recoded at Level III or the picture must be recoded at Level II. Thus, it is possible that in addition to measuring encoding, there may be a small recoding component in the picture-word matching task.

One implication of the present analysis is that both production tasks involve encoding and decoding and that the comprehension task involves encoding and possibly recoding. This observation is important because as some researchers (e.g., Dornic, 1979; Magiste, 1979) have classified production tasks as measures of decoding and comprehension tasks as measures of encoding. It appears, however, that the only valid distinction that can be made is that the decoding component is comparatively larger in production than comprehension tasks. For the remainder of the paper, the tasks are refered as either measures of production or comprehension. An exception is made in discussion of the category matching task. Deciding whether two words are from the same category requires semantic (Level III) processing of both words and thus reflects the time to encode two words and decide if they are from the same category.

Bilingual Balance

Most researchers make the distinction between balanced and nonbalanced bilinguals. (e.g., Dornic, 1979; Ehri & Ryan, 1980; Lambert, 1955). Individuals who demonstrate equal proficiency in both languages on some measure of linguistic fluency or competency are classified as balanced bilinguals. Individuals who demonstrate superior performance in one of their two languages are classified as nonbalanced bilinguals, or as dominant in the language of superior performance (e.g., English-dominant).

A variety of performance measures, such as picture and word naming, have been used to classify subjects as either balanced or nonbalanced bilinguals. The rationale for developing performance indices was first articulated by Lambert (1955) who argued that "automatic behavior is characterized chiefly by its speed" (p. 197). He found that the relative speed with which subjects responded to simple directions in French and English correlated with their initial classification as either balanced bilinguals or English-dominant speakers, based upon language acquisition history and use. Balanced bilinguals evidenced equivalent French-English latencies, whereas English-dominant speakers responded faster in English than French.

Each of the comprehension and production tasks used in the present study was developed originally as a measure of linguistic balance. In this context, the question of interest was whether there was a significant difference between first and second language latencies. Magiste (1979) was one of the first

researchers to demonstrate that such tasks could be used to study the acquisition of bilingual comprehension and production skills at an elementary level of information processing. She accomplished this by comparing actual response latencies of balanced bilinguals, nonbalanced bilinguals and monolinguals. Thus, Magiste (1979) was able to use these same tasks to address two very different questions: "How much time does it take to obtain a comparable competence in two languages and how is response speed influenced by the coexistence of two languages in the same individual?" (p. 79). Before describing Magiste's findings, the development of picture-naming, word-naming, and picture-word matching tasks, as measures of bilingual balance, is reviewed.

Naming pictures in two or more languages is one of the most commonly used measures of linguistic balance (Albert & Obler, 1978; Paivio & Begg, 1978). Ervin (1961a) developed a language dominance test comprised of naming 120 pictures of simple objects in English and Italian. She reported that Italian-dominant adults (who came to the United States after age nine and spoke Italian most of the time) had faster naming latencies in Italian than English, while English-dominant subjects (who came to the United States before age six and spoke English in school and with friends) were faster in English than Italian. Similar results were obtained when a similar test was administered to a group of Navaho-English subjects (Ervin, 1961b). Since its introduction, numerous researchers have used various forms of the picture-naming task (Albert & Obler, 1978; Dornic, 1979; Goodman

et al., 1985; Magiste, 1979; Potter et al., 1984).

One of the first studies using a word-naming task with bilinguals was done by Lambert, Havelka, and Gardner (1959). They hypothesized that the length of the recognition threshold for reporting words in both languages would be a good index of degree of bilingualism. French-English bilingual adults viewed French and English words in a tachistiscope. Word-naming thresholds correlated .46 with the pre-established categories of balanced and nonbalanced bilingualism, based upon language acquisition history. In later studies word-naming latencies, rather than recognition thresholds, were the dependent measures. On the basis of language acquisition history and use, Kolers (1966a) classified subjects as either balanced or nonbalanced bilinguals. He reported that subjects who were more fluent in their first language read sentences in that language significantly faster than in their second language. Similar findings were reported when subjects were asked to read colour names aloud (e.g., Nott & Lambert, 1969; Preston & Lambert, 1969) or read aloud high and low frequency words (Magiste, 1979, 1982).

Measurement of linguistic balance using comprehension tasks has not received as much research attention as measurement with production tasks (Dornic, 1979). In the work that has been done, a variety of complex tasks have been used. For example, subjects have been asked to read a set of directions and perform a motoric response (e.g., Dornic, 1979; Lambert, 1955; Magiste, 1979), and to search for a concept in text (Dornic, 1977). Differences between the time to perform the tasks in each language were

reported to be a sensitive index of linguistic balance. It should be noted that in many of these tasks, subjects had to encode complete sentences, thereby escalating the level of task difficulty beyond the one-word level.

Macnamara (1967) studied comprehension skills at the one-word level in English-dominant adults. Simple line drawings of common objects with a word printed below were presented. Subjects were asked to indicate whether the word and picture matched by pressing one of two response keys. Mean cumulative time for the 18 French words (1230 ms) was significantly longer than for the English words (1089 ms). Goodman et al. (1985) also used a picture-word matching task and reported that number of errors was a useful criterion for classifying children's degree of linguistic balance. Thus, it appears that when working in their second language, nonbalanced bilinguals respond more slowly and make more errors than balanced bilinguals.

Effects of Exposure

As mentioned, Magiste (1979) was one of the first researchers to explore the pattern of comprehension and production latencies as a function of the number of years an individual was exposed to a second language and to compare bilinguals with monolinguals. A cross-sectional method was used to study the effects of years of exposure. Her sample included German speakers, 13 to 18 years of age who lived in Sweden and who were being instructed in German and Swedish at school. In this program, students began reading German, but by grade 2 they

were also taught reading and writing in Swedish. However, not all subjects began the program in grade 1. Thus, Magiste was able to divide the sample of German speakers into eight groups of same-aged subjects, ranging from "0 to 1" to "15 to 17" years of exposure to Swedish. The comparison group was comprised of Swedish monolinguals, 14 to 17 years of age. All subjects were tested on several different comprehension and production tasks, including: picture naming (75 pictures of common objects), word naming (40 words in each language, four to seven letters in length), and eight simple encoding tasks in each language that placed demands on elementary language processing (e.g., "Mark the third letter from the left").

As expected, Magiste (1979) found that latencies in the second language decreased as the number of years of exposure increased. What was surprising was that she found an increase in first language latencies corresponding to the number of years of exposure to the second language. Also, she noted that subjects who had reached a point of bilingual balance took significantly longer to work in either language than monolinguals. For example, balanced bilinguals took on the average 1200 ms to name a picture in either language, whereas Swedish monolinguals took only 850 ms, a difference of 350 ms. This result was not an artifact of original differences between the two groups, as German-dominant speakers, with less than one year of exposure to Swedish, also named pictures in 850 ms in German. The pattern was identical for the word-naming data (monolinguals took 370 ms and balanced bilinguals took 500 ms) and the encoding data

(monolinguals took 280 ms and bilinguals took 350 ms).

Similar findings were reported for a group of balanced and nonbalanced English-German and German-Swedish bilinguals (Magiste, 1982). Balanced bilinguals took longer than nonbalanced bilinguals to name pictures of high and low frequency words in their first language. The increase in first language latencies reported for balanced bilinguals was more pronounced for low than high frequency words.

Reduced automaticity and interference between the two languages (interlingual interference) were hypothesized to account for this pattern of latency scores. Magiste (1979, 1982) based her argument on Shiffrin and Schneider's (1977) model of automatic and controlled processing. An operation is considered to be automatic if, as a result of extensive practice and exposure, it can be performed very quickly and with minimal attentional resources. Controlled operations are much slower and require considerable attention (Shiffrin, Dumais & Schneider, 1981). Magiste (1982) argued that bilinguals' responses were slower in both languages because they had to practice in both linguistic codes. Thus, the total amount of practice in each language was substantially less than the amount of practice monolinguals had in one language. Differences would also be expected between picture and word naming. Given the same exposure to a French word and picture, word naming may be more automatized because it is a case of direct mapping between one word and one verbal label, whereas picture naming is a case of varied mapping between one picture and two verbal labels.

With respect to interlingual interference, Magiste reasoned that a monolingual had only one available verbal label to attach to a concept, whereas a bilingual had two. Thus, when asked to name a picture or decide whether a picture and a word matched, the bilingual activates two verbal labels and then has the additional task of choosing between two competing responses. This choice places demands on controlled processing, thereby increasing response latencies (Magiste, 1982). With respect to word-naming, the bilingual does not have to activate any labels, but rather has to choose between two competing phonological codes, a factor that has been found to increase processing time (Albert & Obler, 1978; O'Neill & Huot, 1984).

The main conclusion from Magiste's (1979, 1982, 1984) research is that being equally competent in two languages increases comprehension and production latencies in each language relative to nonbalanced bilinguals and monolinguals. Magiste studied elementary operations but, "it is these simple processes that most often cause a 'chain reaction'—a considerable deterioration in more complex operations" (Dornic, 1979, p. 331). However, it is premature to draw implications from Magiste's research to Canadian anglophones enrolled in French programs because there are several issues that warrant further examination, including: (a) the replicability of Magiste's results in the Canadian linguistic context when subjects differ with respect to intensity of exposure to French; (b) examination of the effects of duration of exposure and age; (c) examination of the amount of variability in comprehension and production

latencies associated with one individual difference variable, intelligence level (see Verbal Intelligence); and (d) examination of the relative rate of development of comprehension and production when potential sources of confounding in the tasks, such as level of task difficulty and stimulus items, are controlled (see Differential Rates of Skill Development).

First and foremost is the issue of whether Magiste's (1979) results can be replicated with Canadian anglophone students when intensity of exposure varies from intense (immersion program) to minimal (extended French program). Two studies relate to this issue, one conducted in the United States (Potter et al., 1984) and one in Canada (Favreau & Segalowitz, 1982). Potter et al. studied the picture-naming skills of two groups of American bilinguals: adult Chinese-English balanced bilinguals, and English-dominant high school students who had studied French in school for two to three years. Among adult subjects, latencies to name pictures in the first and second language were more similar (910 and 1050 ms, respectively) than those reported for the adolescents (850 and 1250 ms, respectively). Furthermore, balanced bilinguals took longer to name pictures in their first language than did nonbalanced bilinguals (910 and 850 ms, respectively).

Magiste's (1979) findings regarding word-naming skills were addressed indirectly in a study by Favreau and Segalowitz (1982), who examined the reading and listening skills of Canadian French-English adult bilinguals. Their classification of subjects differed from the usual distinction between balanced and

nonbalanced bilinguals because they classified subjects on the basis of optimal reading rate, a measure of oral reading speed and comprehension. Subjects were divided into two groups. In the equal-reading-rate group, the difference between optimal reading rate in each language was less than 10%, whereas in the unequal-reading-rate group the difference was greater than 10%. Subjects in the equal-reading-rate group took approximately the same amount of time to read and listen to text in English and French. In contrast, in the unequal-reading-rate group, processing was faster in the first than second language. Comparison of processing rates for the two groups revealed that subjects in the unequal-reading-rate group were faster than subjects in the equal-reading-rate group on first language materials but slower on second language materials. Favreau and Segalowitz (1982) commented that this finding was unexpected and concluded that this result, "suggests the possibility of a trade-off when individuals are highly skilled in their two languages" (p. 339).

As previously mentioned, Magiste (1979) did not evaluate the relationship between effects associated with duration of exposure and age. For each exposure category the mean age of subjects was equivalent, although the range in age was from 13 to 18 years. Given that a decrement in picture- and word-naming latencies is usually associated with age (e.g., Goodman et al., 1984; Hatch, 1978), it is possible that the reported increase in first language latencies is an underestimate and the decrease in second language latencies is an overestimate. By including

immersion and extended French pupils at two grades, it is possible to explore the relationship between effects associated with age and duration of exposure (comparison across Grade), effects associated with intensity of exposure (comparison across Program), and the relationship between intensity and duration of exposure. Several specific questions were investigated.

Question 1. Do immersion and extended French pupils show decreases in French latencies across grade that are similar in magnitude?

Question 2. Do immersion pupils respond faster in French than extended French pupils at the same grade?

Question 3. Do grade 9 immersion and extended French pupils take longer than their grade 6 counterparts to respond in English or does the expected age-related decrease in response latencies obscure increases in latencies expected for duration of exposure?

Question 4. Do immersion pupils take longer to respond in English than extended French pupils in the same grade?

Verbal Intelligence

Magiste (1979, 1982) did not evaluate the extent to which her findings may have been attributable to differences in intelligence. There have been numerous studies examining the effect of bilingualism on intellectual ability (e.g., Ben-Zeev, 1977; Bruck, Lambert, & Tucker, 1974; Ianco-Worrall, 1972; Oren, 1981; Peal & Lambert, 1962), but the effect of intelligence on the acquisition of skills in a second language has received

little attention (McLaughlin, 1977). One exception is the research on Canadian French immersion pupils. In a study of grade 11 French immersion pupils (Genesee, Allister, & Morin, 1974), IQ was correlated positively with performance on formal and written aspects of French language skills (e.g., tests of grammatical knowledge) but not with listening and speaking skills. In a review paper summarizing the data from students in immersion programs from kindergarten to grade 9, Genesee (1976) stated that the acquisition of literacy skills in French is associated with the student's IQ level, but that the acquisition of interpersonal communication skills is not. A similar conclusion was reached by Swain and Lapkin (1981). "The ability to communicate functionally in the second language is not related to intelligence as measured by Standardized IQ tests" (p. 129).

In a recent study aimed at finding measures that differentiated between students who performed well in immersion programs and those who dropped out, Trites (1981) reported that the average Verbal IQ score for all kindergarten immersion subjects was 118. The average Verbal IQ score of the students remaining in the immersion program at grade 1 was 124. Among those students who had dropped out, the average Verbal IQ was 115. Although Trites reported that there were other variables associated with the group of students who droppped out, it appears that Verbal IQ is a variable worthy of further investigation.

Based on the results of these studies, it appears that Verbal IQ may be related to acquisition of comprehension and

production skills in a second language. Therefore, Verbal IQ was included in the present study as an additional modification to Magiste's (1979) design. Given that the examination of Verbal IQ was exploratory in nature, subjects were not chosen to form different Verbal IQ groups. Rather, the Verbal IQ score for each subject was recorded and relationships between Verbal IQ and the different performance measures were examined.

Question 5. Is there a relationship between Verbal IQ and comprehension and production latencies in both languages?

Differential Rates of Skill Development

One final issue of concern involves Magiste's finding that a point of balance was reached on word naming (a measure of production) after only four years of exposure, whereas balance was achieved on the comprehension task after five years of exposure. Interest in whether production skills develop prior to comprehension skills is not new. As early as 1963, Kolers questioned whether comprehension and production skills were similar processes that ran off in reverse order. He concluded from his examination of bilinguals' reading and speaking skills that they were not symmetrical skills. From their review of the literature, Albert and Obler (1978) reached a similar conclusion. They noted that although a positive correlation had been found between comprehension and production measures (e.g., Lambert, 1955), comprehension skills in the second language often were reported to precede production skills. In studies of anglophone Canadian students in French immersion programs, it has been found

that by grade 8 students perform below native-like competency levels on complex production tasks in French (Harley & Swain, 1977, 1978; Szamosi, Swain & Lapkin, 1979), but on comprehension tasks performance levels are equivalent to bilingual francophone pupils (Swain & Lapkin, 1981).

There are several problems inherent in this body of research, including: (a) the use of different stimulus items in comprehension and production tasks with little attention paid to equating items in terms of frequency ratings; and (b) measurement of different aspects of linguistic knowledge in production and comprehension tasks. Magiste's (1979) study illustrates this latter criticism. Both production tasks she used, picture and word naming, assessed knowledge at the one-word level, but in the comprehension task (e.g. "Mark the third letter from the left) subjects had to have a basic knowledge of grammar in order to comprehend the instructions. Thus, the finding that word-naming skills developed more rapidly than comprehension skills may have been due to differing levels of task difficulty.

Magiste (1979) reported that bilingual balance was achieved earlier on word naming (five years) than on picture naming (six years). In both tasks, single-item stimuli were used. Thus, potential confounding due to differing levels of task difficulty was not a problem. However, given that the frequency ratings of the items in each task were not reported, the results may have been confounded by frequency differences.

In the present study, the comprehension and production tasks were designed to circumvent the aforementioned problems by

using the same 48 items for all three tasks and by choosing comprehension and production tasks that assessed skills at the one-word level. The same 48 items were used instead of counterbalancing materials because there was a limited number of French nouns that grade 6 extended French pupils knew.

Question 6. When stimulus materials and level of task difficulty are equated, will bilingual balance be achieved first on picture naming, word naming, or picture-word matching?

Facilitation Effects

Using the same stimulus items introduces the possibility of facilitation effects. Cross-language facilitation may result when there is exposure to the stimuli first in one language and then the other. Cross-task facilitation may occur when items are presented first in a comprehension task and then in a production task or vice-versa. Practice effects may occur when the same words are repeated within a task in the same language. Although examination of facilitation effects is only a minor focus in this study, each type has different implications for second language processing.

If cross-language facilitation effects are found, the implication is that presentation of a word in one language primes response to the same concept in the other language. Among balanced bilinguals, cross-language or interlingual facilitation effects are of the same magnitude regardless of the language in which the word first appeared (Kolers, 1966a; Kolers & Gonzalez,

1980). Among nonbalanced bilinguals, Ehri and Ryan (1980) found a significant facilitation effect in the nondominant language when words were presented first in the dominant language but no corresponding faciliation effect when the order of presentation was reversed. Ehri and Ryan's findings may indicate that processing in the second language is too slow to facilitate processing in the first language. In the present study, the direction of cross-language facilitation effects was determined by examining whether facilitation was greater when words were presented first in French and then in English (Test Sequence French-English) or first in English and then in French (Test Sequence English-French).

Question 7. Does receiving the picture, word, or picture-word pair first in English facilitate performance in French more than the converse?

Investigation of cross-task facilitation effects was exploratory. Interest in this issue was generated from the intuitive assumption that if a subject received the comprehension task (picture-word matching) prior to the production task (picture naming), then latencies to name pictures would be smaller than when picture-naming preceded picture-word matching.

Question 8. Does receiving the picture-word matching task first facilitate performance on the picture-naming task more than the converse?

In the picture-word matching task, each word was repeated twice in the same language and thus a practice effect was expected. It has been found that practice effects are larger in

the nondominant than dominant language (Kolers, 1966a; Macnamara, 1967).

Question 9. Will practice effects be larger in French than English for English-dominant students?

B. Organization of Bilingual Semantic Memory

As previously mentioned, one of the unique aspects of the present study is that the examination of comprehension and production skills is extended to include examination of how second language learners organized semantic knowledge in memory. Of particular interest is whether there is a relationship between level of comprehension (encoding) skill in each language and type of structural organization in memory. In order to explain why this is an important issue, it is necessary to review the main developments in the research on bilingual semantic memory.

Two opposing models of bilingual semantic memory have been developed. Proponents of the independent or two-store model postulate that bilinguals have two separate memory systems, one for each language. Information in each system is not necessarily available without recourse to some special procedures of translation (e.g., Kintsch, 1970; Kolers & Gonzalez, 1980; Tulving & Colotla, 1970). Proponents of the interdependence, or one-store model, hypothesize that bilinguals have one node for each concept to which verbal labels for each language are attached. Within this model, no mechanism is required to provide communication between the languages (e.g., McCormack, 1977; Potter et al., 184; Shannon, 1982; Snodgrass, 1984).

A variety of paradigms have been used to investigate the bilingual memory storage issue. Evidence supporting a two-store model has been found in studies of recognition memory (e.g., Kintsch, 1970), proactive inhibition (Dillon, McCormack, Petrusic, Cook & Lafleur, 1973; Goggin & Wickens, 1971), interlingual word association (Kolers, 1963a), free recall with attenuated lag effect (Glanzer & Duarte, 1971), and switching time (Macnamara, 1967; Macnamara, Krauthamer & Bolgar, 1968; Macnamara & Kushnir, 1971) although two recent studies refute the claim that bilinguals can switch off processing words in one language when working in the other (Goodman et al., 1985; Guttentag, Haith, Goodman, & Hauch, 1984). In support of a one-store model, it has been found that (a) translation of a word facilitates recall as well as repetition of the word (Kolers, 1966b), (b) on bilingual versions of the Stroop task subjects cannot ignore words printed in the other language (Dryer, 1971; Magiste, 1984; Preston & Lambert, 1968), and (c) among balanced bilinguals interlingual interference is equivalent to intralingual interference on a picture-word interference task (e.g., Ehri & Ryan, 1980).

Faced with these discrepant findings in the literature, Snodgrass (1984) noticed that results are consistent with a two-store model when tasks eliciting episodic memory and a delayed response (e.g., free recall) are used. When tasks eliciting semantic memory and on-line responding were employed, results support a one-store model. In using the term episodic memory, Snodgrass was referring to memory that was idiosyncratic

and contained information about the occurrence of events in time and space. Recalling a list of unassociated words would be an example of a task assessing episodic memory, whereas deciding whether peach and apple were from the same category would be an example of a task assessing semantic memory. Snodgrass questioned whether subjects might not only remember which concept was presented (i.e., language-independent information), but also what language it was in (i.e., language-dependent information). The point that Snodgrass seems to be making is that when the task requires subjects to process words to the semantic level, they focus on meaning and not which language the words are in. As a result the data from these studies support a one-store model. Alternatively, in a task such as recall of a word list, subjects differentiate between the words in terms of the language of presentation. Thus, when subjects work at this lexical level, it is to their advantage to store word knowledge for each language separately. Shannon (1982) noted also that subjects could code words in terms of language specific cues (i.e., language-dependent) or general semantic cues (i.e., language-independent) and concluded that perhaps the question of bilingual memory storage was not categorical but depended on the demands of the task.

Several authors have attempted to translate this theoretical distinction between language-dependent and language-independent knowledge into a model of bilingual language processing (Guenther, 1984; Potter et al., 1984; Shannon, 1982). Guenther (1984) postulated that lexical processing, similar to

Snodgrass's (1984) Level II processing, is language-dependent, whereas semantic processing (Level III) is language-independent. Thus, tasks that elicit on-line semantic processing would support a one-store model and tasks that elicit episodic memory would support a two-store model.

Given the aforementioned distinction, it seems that one of the better ways to study bilingual memory organization would be to use a task known to elicit on-line semantic processing, such as a semantic verification or category matching task. Caramazza and Brones (1980) used a semantic verification task in their study of adult balanced Spanish-English bilinguals. Subjects were requested to press one of two buttons to indicate whether the exemplar matched the category, irrespective of whether the category and exemplar were presented in the same or different languages (e.g., fruit-pomme, match; fruit-chair, mismatch). In their first experiment, there was a two-second interval between presentation of the category and exemplar and in the second experiment the words were presented simultaneously. Analyses of both experiments revealed no significant difference between response latencies in the same versus different language conditions.

The Potter et al. (1984) study of Chinese-English bilinguals differed methodologically from that of Caramazza and Brones (1980) in several respects, including auditory presentation of the category name, a 500-ms delay between the category label and visual presentation of the exemplar, and the requirement of a "yes" or "no" response. In spite of these

methodological differences, Potter et al.'s (1984) results were similar to Caramazza and Brones (1980) earlier findings in that no significant differences were found between the same and different language conditions. Both studies indicate that in the different language condition, bilingual subjects do not have to translate between two separate memory stores. Rather, they work from one semantic memory store.

To recount, several authors have put forward the argument that the original statement of the bilingual semantic memory controversy needs to be reformulated to take into account findings that information may be stored in either a language-dependent or language-independent manner. Another reformulation of the original model has been proposed concerning whether the organization of bilingual semantic memory undergoes structural change as competency in the second language changes.

Gekoski (1980) proposed that individuals with little second language proficiency have two memory stores and translate through the first one when working in the second. Expanding upon Gekoski's (1980) structural change hypothesis, Potter et al. (1984) proposed that second language learners start out with only lexical associations between two memory stores. Slowly, direct links develop between the second language lexicon and concepts. Finally, when bilingual balance is achieved there is one common memory store. In this single memory store, each concept has verbal labels from both languages attached to it.

The structural change hypothesis was addressed by Potter et al. (1984), who compared adult balanced and adolescent

nonbalanced bilinguals. Subjects were asked to name pictures in both languages, name words in one language, and translate words from one language to the other. The time to translate a Chinese word into English did not differ significantly from the time to name a picture in English, indicating that nonbalanced and balanced bilinguals utilized one common memory store and did not translate via direct associations between two separate memory stores. Similar conclusions can be drawn from Goodman et al.,'s (1985) study, although the task (picture-word interference) and dependent measures (interlingual and intralingual interference) were different. These authors compared young nonbalanced ($\underline{M}$ = 6.2 years) and older balanced ($\underline{M}$ = 10.2 years) bilinguals. They reported that the pattern of intralingual and interlingual interference was virtually identical for all groups of subjects and concluded that, "the second language words make direct contact with meaning soon after they can be identified" (p. 116).

From the results of the aforementioned studies, structural organization of bilingual semantic memory does not appear to change as a function of the level of second language proficiency, but differences in the tasks utilized makes it difficult to draw any definitive conclusions. In the present study, a category matching task was used. This made it possible to compare the present results to previous studies of adult bilinguals. In line with previous research, latencies on same and different language trials were compared to determine whether storage of semantic knowledge was most consistent with a one- or two-store model of memory. Support for a two-store model is indicated if latencies

for different language trials are longer than for same language trials by some constant, which may be assumed to represent translation time. Alternatively, support of a one-store model is indicated if latencies are equivalent on same and different language trials.

One problem with previous studies of bilingual semantic memory is that no objective measure of balance, such as encoding latencies, was employed. Rather, subjects were classified as balanced or nonbalanced bilinguals on the basis of language acquisition history and self-rating scales (e.g., Caramazza & Brones, 1980; Potter et al., 1984). Given that it is assumed in the structural change hypothesis that subjects with equivalent latencies (i.e., balanced bilinguals) utilized one memory store, it seems necessary to include such a measure of balance. In the present study, inclusion of French-French and English-English word pairs afforded a measure of encoding speed in each language. By comparing French-French and English-English latencies, it can be determined whether encoding speeds are equal in each language, or faster in one of the two languages. Consideration of the possible relations between the two encoding speeds (equal versus different), and the two types of memory systems (one- versus two-stores) made it possible to generate four models of second language processing and storage, in place of the original two. The models are: One-Store Different Encoding, One-Store Equal Encoding, Two-Store Different Encoding, and Two-Store Equal Encoding.

Although this reformulation of the original bilingual semantic memory models may appear to be a major departure from previous research, it parallels developments in the picture-word processing and storage research (Glucksberg, 1984; Potter et al., 1984; Snodgrass, 1984). In both areas, the issues are: (a) whether information is stored in a modality independent (one-store) or modality dependent (two-store) memory structure; and (b) whether information from each modality (language) is encoded at the same or different speeds when processing occurs at a semantic level. For example, Pellegrino, Rosinski, and Siegel (1977) developed a model to study whether words and pictures were stored separately or together and whether encoding speeds for words and pictures were equivalent when categorical decisions were required. Given the similarity between the issues in the picture-word and bilingual semantic memory controversies, Pellegrino et al.'s (1977) model was adapted for the present investigation of bilingual semantic processing and storage.

Description of the Category Matching Task

A category matching task, rather than a semantic verification task, was chosen because children in the younger grade were not as familiar with category labels and the aim was to avoid the confounding effects of task difficulty. High-frequency words from two categories, body parts and animals, were chosen. Two words, such as arm and orange, were presented simultaneously. Students were asked to decide whether the two words were from the same category (match) or from two different

categories (mismatch) and to press on the appropriate response button. The two words were either in the same language (English-English or French-French) or were from two different languages (English-French or French-English), for a total of four language conditions. The only difference in the latter two conditions was the relative position (i.e., to the right or left of the midline) of the French and English words. Subjects were instructed to make their decision about category membership irrespective of the presented language.

The one- and two-store models are illustrated in Figure 2. Separate parameters, representing encoding speed in English (ae) and French (af) are included in both models. As mentioned, in both models encoding speeds in French and English can be either equal or different. In the Equal Encoding models, ae and af are assumed to be equivalent. In the Different Encoding models, af is assumed to be longer than ae. Encoding speed is assumed to be longer in French because all subjects learned English as their first language and spoke it at home. The translation parameter, t, differentiates between one- and two-store models. In the One-Store models, all words are assumed to be stored in one memory structure. Thus, when given a pair of words in which one is in French and the other in English, the subject can access both word meanings without having to translate. In the Two-Store models, it is assumed that translation is required on the different language trials. The parameter, d appears in both models and represents the time to decide if two words are from the same category and respond. In the present study, d cannot be

estimated separately and thus no comparison was developed to test for this parameter.

Insert Figure 2 about here

Referring to Table 1, three orthogonal comparisons, (a) through (c), were developed to discriminate between the four models. The first comparison involves differences between English-English and French-French trials and addresses whether encoding speed is equal or different in the two languages. The second comparison tests differences between same and different language trials. If the difference between same and different language trials is nonsignificant, then a one-store model is consistent with the data. However, if latencies for different language trials are longer than same language trials by some constant, assumed to represent translation, then a two-store model is more consistent with the data. The third comparison deals with differences between the two different language trials, French-English and English-French, and is a test for possible asymmetry effects in the direction of translation in models that involve a translation parameter.

Insert Table 1 about here

The One-Store Different Encoding Model represents a case in which the difference in processing French and English words lies in the speed of encoding. It is one possible form of nonbalanced

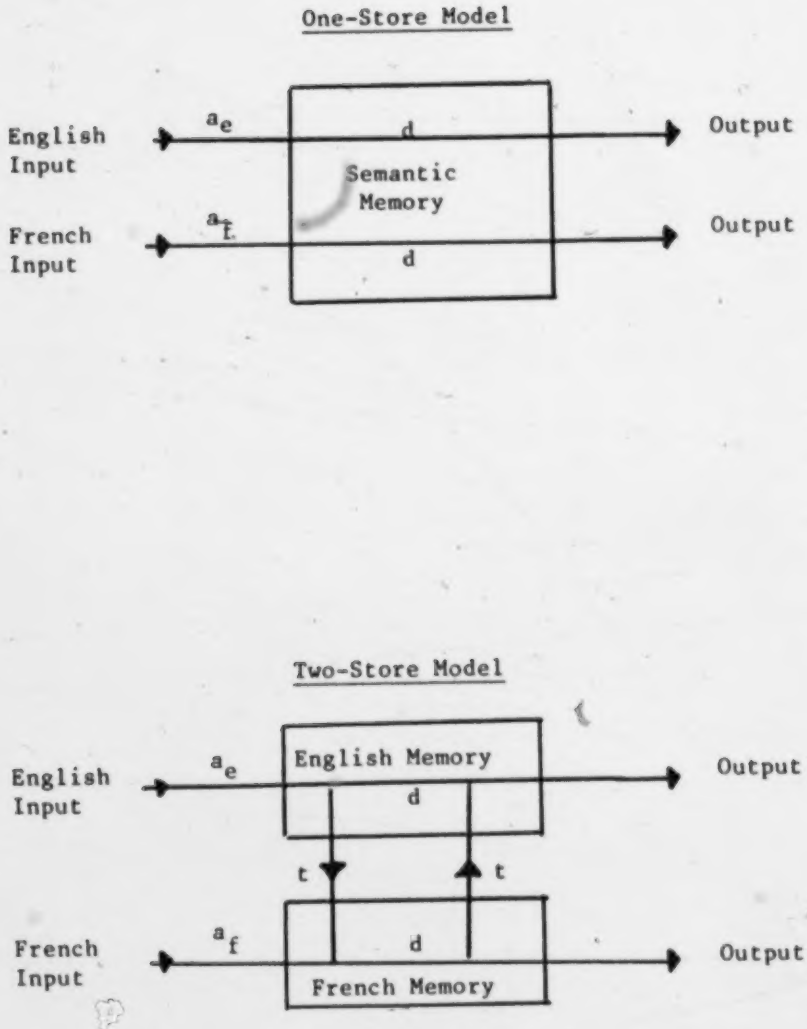


Figure 2. One-Store and Two-Store models of bilingual semantic memory. (Note: a_e = English encoding time; a_f = French encoding time; t = translation time; d = time to decide if two words are from the same category)

Table 1 Predictions for Latency Data From the Category Matching Task

Language Condition	One-Store Different Encoding	One-Store Equal Encoding	Two-Store Different Encoding	Two-Store Equal Encoding
French-French (FF)	$2a_f + d$	$2a + d$	$2a_f + d$	$2a + d$
English-English (EE)	$2a_e + d$	$2a + d$	$2a_e + d$	$2a + d$
French-English (FE)	$a_f + a_e + d$	$2a + d$	$a_f + a_e + d + t$	$2a + d + t$
English-French (EF)	$a_e + a_f + d$	$2a + d$	$a_e + a_f + d + t$	$aa + d + t$
	Predicted Results for Orthogonal Multiple Comparisons			
(a) FF-EE	FF > EE	0	FF > EE	0
(b) $\frac{EF+FE}{2} - \frac{FF+EE}{2}$	0	0	$\frac{EF+FE}{2} > \frac{FF+EE}{2}$	$\frac{EF+FE}{2} > \frac{FF+EE}{2}$
(c) FE-EF	0	0	FE ? EF	FE ? EF

Note. a_e = parameter representing time to encode and retrieve the meaning of an English word; a_f = parameter representing time to encode and retrieve the meaning of a French word; d = parameter representing time to make a category membership decision, but it cannot be estimated separately from either a_e or a_f; t = parameter representing the time to transfer knowledge from one memory store to the other.

bilingualism. With respect to the first comparision (Table 1), encoding on French-French trials is assumed to take longer than on English-English trials. Second, latencies for same and different language trials are assumed to be equal because translation is not involved. Lastly, because translation does not occur, English-French latencies are assumed to be equivalent to French-English latencies.

The One-Store Equal Encoding Model represents a case in which there is no difference in processing in any of the four language conditions. This is one possible form of balanced bilingualism. With respect to the first comparison, latencies on French-French and English-English trials are assumed to be equivalent. The expected findings for comparisons (b) and (c) are identical to those for the One-Store Different Encoding Model as in both models translation is not involved.

The Two-Store Different Encoding Model represents a case in which processing English and French words involves a difference in encoding speed and translation. It is another possible form of nonbalanced bilingualism. First, encoding latencies are assumed to be longer on French-French than English-English trials. Second, latencies for different language trials are assumed to be longer than same language trials by some constant, representing translation time. In the third comparison, possible asymmetry in the direction of translation is indicated, but no assumption is made regarding whether latencies are longer on French-English or English-French trials. The rationale for assuming there may be asymmetry effects is based on research

indicating that translation of words from the nondominant to dominant language is faster than the converse (Gekoski, 1981; Taylor, 1971). In the present task, it cannot be determined with any degree of certainty in which direction subjects translate for a given pair of words, such as French-English. One possible hypothesis is that the language of comparison is determined by the first word (i.e., to the left of the midline). If this is the case, and if there is asymmetry in the direction of translation, then a difference would be expected between French-English and English-French word pairs. A minimal test of asymmetry in $\underline{t}$ is the third comparison listed for both two-store models.

The Two-Store Equal Encoding Model, represents a case in which processing French and English words involves translation. It is one possible form of balanced bilingualism. Latencies are assumed to be equal on French-French and English-English trials, as indicated in the first comparison. Second, different language trials are assumed to be longer than same language trials by some constant representing translation. For the reasons previously mentioned, possible asymmetry in $\underline{t}$ was tested in the third comparison.

C. Summary of the Issues

Examination of the relevant literature on bilingual comprehension, production and organization of semantic memory reveals that developmental issues have received scant attention, particularly with respect to anglophone Canadian students

learning to speak French as a second language. Furthermore, there is virtually no research examining the relationship between speed of encoding in each language and organization of semantic memory.

Four groups of second language learners, grade 6 and 9 immersion and extended French pupils, were assessed on two production tasks, one comprehension task, and one task measuring organization of semantic memory. Thus, it was possible to explore the effects of duration and intensity of exposure to the second language on the development of comprehension, production, and memory organization.

The first four questions focus on the effects of duration and intensity of exposure on the development of comprehension and production skills. In the fifth question, the issue is raised as to whether a relationship will be found between Verbal IQ and word-naming, picture-naming and picture-word matching latencies. Question six deals with whether a point of balance will be achieved earlier on picture-word matching, word naming, or picture naming when stimulus materials and level of task difficulty are equated. Issues relating to facilitation effects are mentioned in questions seven to nine. The questions developed for the category matching task focus on whether subjects in each of the four groups will utilize one- or two-memory stores and whether the relative speed of encoding in French and English will be related to the structural organization of semantic memory.

II. METHOD

A. Subjects

Ninety-six anglophone students participated in the study, including 12 males and 12 females in each of four groups: French immersion, grade 6; French immersion, grade 9; extended French, grade 6; and extended French, grade 9. Participation in the study was voluntary and those who chose to participate were paid five dollars at the end of the second session.

Immersion students had been instructed only in French from kindergarten until grade 2, at which time English was introduced for one subject and gradually increased to include about 40% of the curriculum by grade 9. All of the grade 6 immersion and 79% of the grade 9 immersion pupils began the program in grade 1. The remaining grade 9 immersion pupils began the program in grade 2.

Extended French students had been instructed only in English from kindergarten until grade 4, at which time they received four hours of French instruction per week. The amount of French instruction was gradually increased from grades 4 to 9 to include one subject taught in French (e.g., social studies). All grade 6 and 9 extended French pupils began the program in grade 4.

Median ages (in years:months) for the four groups were 11:11 for grade 6 immersion (range 11:7 to 12:4), 11:9 for grade 6 extended French (range 11:6 to 12:4), 14:7 for grade 9 immersion (range 14:5 to 15:2), and 14:6 for grade 9 extended

French (range 14:4 to 15:1). None of the students tested had a history of a language disorder or learning disability. Verbal intelligence scores from the Canadian Cognitive Abilities Test (Rand, Randhawa, Hunt & Rawlyk, 1974), which had been administered by school personnel six to nine months prior to the study, were available for all but three subjects.

B. Materials and Procedures

Testing sequence and task order. Each subject was tested individually over two sessions. The picture naming, word naming, and picture-word matching tasks were administered in French during one session and in English during the other. The category matching task, involving the simultaneous use of both languages, was administered once at the end of the first session.

There were two Test Sequences for the picture naming, word naming, and picture-word matching tasks: Half of the subjects were tested in English during the first session and in French during the second session (Sequence English-French), and for the remaining students order was reversed (Sequence French-English). Three task orders were created with the constraints that each task appear once as the first task, and that each of the two production tasks (picture and word naming) were preceded by the comprehension task (picture-word matching). This latter stipulation on Task Order was included so that any facilitation effects of the comprehension task on the production tasks could be assessed. For purposes of simplicity, each Task Order is identified by the task that appeared first: P1 refers to the

order picture naming, word naming, picture-word matching; W1 to the order word naming, picture-word matching, picture naming; and PW1 to the order picture-word matching, word naming, picture naming. Testing Sequence, Task Order, and Sex were counterbalanced across subjects within each group.

Stimuli for picture naming, word naming, and picture-word matching. To select items that were likely to be familiar to students at both grade levels, a list was compiled of 200 French nouns that referred to objects and that were taught to immersion pupils in the first and second grade (Faubert, 1979; Ministre d'education de Manitoba, 1980) and then reviewed by two teachers familiar with the extended French and immersion programs. From this list, 9 practice and 48 test words were chosen that met the following criteria: (a) each item was familiar to students of both grade levels; (b) the name of each item was neither a cross-language cognate (i.e., words that are spelled the same but have different meanings in each language, such as the French word pain meaning bread and the English word pain) or confusable visually or acoustically with any other noun within or across languages; (c) names of the items did not exceed 10 letters in either language; and (d) frequency ratings (Kucera & Francis, 1967) for the English words were either above 13 or the norm for age of acquisition (Carroll & White, 1973) was under 3.5 years of age. The same 57 words (i.e., 9 practice and 48 test items) were used for all three tasks. See Appendix A for a list of the nouns and all relevant normative data.

Pictorial representations of the 57 nouns were chosen from Snodgrass and Vandewart's (1980) set of 260 standardized pictures. Name agreement ratings were above 60% (M = 4.01; SD = .7). Name agreement ratings were developed by Snodgrass and Vanderwart (1980), who asked 219 college students to identify 260 pictures by writing down the first name that came to mind upon viewing the picture. The percentage of subjects who used the same word when naming the picture is the index of name agreement. Familiarity ratings which were defined as the degree to which one comes in contact with or thinks about the concept (Snodgrass & Vanderwart, 1980), were above 2.2 (M = 3.9; SD = .67). In developing the familiarity ratings for 260 pictures, Snodgrass and Vanderwart instructed subjects to judge the familiarity of each picture according to how unusual or usual it was in their realm of experience on a five-point scale, with 1 denoting very unfamiliar and 5 denoting very familiar. The mean familiarity ratings were computed for each picture based on the responses of 219 college students. See Appendix B for the pictorial stimuli.

Administration of picture- and word-naming tasks. For each task, the 57 stimuli were mounted on 14 separate sheets of paper. Five pages contained six words or pictures (including one practice page) and nine pages contained three words or pictures (including one practice page). The 14 pages, each 27.5 cm by 21.2 cm, were laminated and made into a booklet. Simple black and white line drawings, 3.5 cm by 2.5 cm, were used in the picture-naming task. Black, bold-face, 8-mm uppercase letters were used in the word-naming task.

Each student was to name aloud the pictures or words from top to bottom and finish by reading the word "Stop" (or "Arret" in the French booklet), printed at the bottom of each page. Practice trials were repeated until the subject demonstrated understanding of the task requirements. On test trials, the subject was instructed to work as quickly as possible without making mistakes.

Response latencies were measured in centiseconds from the moment the students saw the page until he or she said "Stop" or "Arret". A stopwatch was used to record latencies. Errors were recorded. The purpose in obtaining naming latencies for three and six words was to compute the slope of a function relating latency with number of words named. Such a slope would serve as a measure of time to name one word, unconfounded by extraneous time factors that were presumed to be constant, such as saying the word "Stop" and starting and stopping the timer. The intercept of the function would be a constant value reflecting the time for these extraneous factors.

Control of confounding due to order of presentation of the 48 stimuli was achieved by developing four different list orders and by counterbalancing list order within the four task by language conditions. For example, one subject might receive picture naming (English) in list order 1, picture naming (French) in list order 2, word naming (English) in list order 3, and word naming (French) in list order 4.

Administration of picture-word matching. Each picture-word pair was presented on a 15-cm by 10-cm laminated card. Black and

white pictures, 3.5 cm by 2.5 cm, and black, bold-face, 5-mm uppercase letters were used. Each of the 48 words appeared once with a matching picture and once with a mismatching picture, for a total of 96 stimulus cards in French and 96 in English. There were 12 practice cards in each language. Subjects were given the English cards during one testing session and the French card during the other.

Cards were presented in sets of six. Each set had a cover card. Subjects were instructed to hold the cards in one hand, to remove the cover card when the experimenter said "Begin", and to sort the cards into two piles, match and mismatch, as quickly as possible without making mistakes. Testing began following subjects' demonstration of task mastery on the practice trials. Response time was measured with a stopwatch in centiseconds from the moment the cover card was placed on the table until the sixth card was placed down. Errors were also recorded.

Stimuli for the category matching task. Fourteen nouns were chosen from two categories: body parts and animals. These words were part of the original pool of 200 words and met the following criteria: (a) name agreement ratings (Snodgrass & Vanderwart, 1980) were above 60% (M = 91.1; SD = 10.9); (b) familiarity ratings (Snodgrass & Vanderwart, 1980) were greater than 1.8 (M = 3.75; SD = 1.2); and (c) frequency ratings (Kucera & Francis, 1967) were above 13 (M = 63.7; SD = 44) or age of acquisition norms (Carroll & White, 1973) were less than 3.5 years. An additional six words were used to generate the eight practice trials. See Appendix C for normative data on the words

used in the category matching task.

Forty-two true and 49 false word pairs were generated from these 14 words. Then 40 true and 40 false item pairs were randomly selected to form the item pool. From this pool of 80 word pairs, four lists were created, such that each word pair appeared once in all four language conditions. Next, the 80 word pairs within each list were ordered unsystematically in accordance with three constraints: (a) no more than two consecutive trials were from the same language condition; (b) no more than three consecutive trials were either true or false; and (c) no word was repeated in two consecutive trials in either language. See Appendix D for the word pairs. List order was counterbalanced across the four groups.

Administration of the category matching task. Subjects were seated .5 m from a video display monitor. In front of the monitor was a three key response panel which subjects were instructed to work with the index finger of their preferred hand. The left and right response keys were used to indicate if the words matched (same category) or mismatched (different category), and key position was counterbalanced across subjects within each group. Subjects were instructed to press the middle or "Ready" key when they were ready for another word pair. Upon pressing this key, the word "READY" and a fixation point directly below it appeared in the middle of the screen. The word "READY" was accompanied by a brief auditory signal. There was an interval of 10 ms between pressing the ready key and the appearance of the two words (printed in upper case letters) situated to the right

and left of the fixation point, with an average horizontal distance of 4 cm. Upon pressing either of the response keys, the words disappeared and an auditory signal indicating if the response was correct or incorrect occurred immediately. Response latencies were recorded by the computer to the nearest millisecond from stimulus onset until a response key was pressed. Error data were also recorded.

The training procedure began by familiarizing subjects with the response keys. Eight trials were given in which subjects were instructed to press the match key if the word "YES" appeared on the screen, and the mismatch key if the word "NO" appeared. When a criterion of 7 out of 8 correct trials was attained, the practice word trials were administered. Practice word trials, which were identical in format to the test trials, were presented in blocks of 8, until a criterion of 7 out of 8 correct was reached. At this point the 80 test trials were administered.

III. RESULTS AND DISCUSSION

A. Verbal Intelligence Scores

The Verbal IQ scores of immersion and extended French pupils were approximately equivalent at each grade level. However, as shown in Table 2, mean Verbal IQ was higher among grade 9 pupils ($\underline{M}$ = 127.2) than grade 6 pupils ($\underline{M}$ = 117.2). If Verbal IQ is related to task performance, then any main effects or interactions involving grade would be confounded.

Insert Table 2 about here

Two steps were taken to deal with this problem. A series of correlations were computed to determine whether Verbal IQ was related to performance measures for any of the four tasks. For those tasks where significant correlations were found, analyses of covariance were computed rather than analyses of variance. Analysis of covariance was used to statistically remove from the dependent variable any variance attributable to the covariate. However, the assumption of homogeneity of regression cofficients must be met if the results are to be considered valid. That is, the form of the relationship between the covariate and dependent variable must be linear and the same for each treatment group (Keppel, 1973). This assumption can be tested by examining the covariate by treatment interaction. If the interaction is nonsignificant, then it can be assumed that the effect of the covariate is uniform for all groups. This procedure for

Table 2

Mean Verbal IQ Scores for French Immersion and Extended French Pupils

Group	n[a]	M	SD
FI6	24	118.7	13.4
EF6	24	116.7	11.5
FI9	24	124.5	12.0
EF9	21	129.8	6.7

Note. FI6 = French immersion, grade 6; EF6 = Extended French, grade 6; FI9 = French immersion, grade 9; EF9 = Extended French, grade 9. The same notation is used in all following tables.

[a]Number of students for whom a Verbal IQ score was available.

examining the homogeneity of regression coefficients was followed whenever necessary.

For the picture-naming, word-naming, and picture-word matching tasks, correlations were computed between Verbal IQ scores and French, English, and French-English differential latencies. Correlations were computed between absolute French and English latency scores and differential latencies because it has been reported in the literature that both are useful indices of linguistic skill. Traditionally, absolute latencies have been used as the measure of skill in each language (e.g., Ervin, 1961a, 1961b; Lambert, 1955). Recent research findings have indicated, however, that it is the relative skill level in each language that is the most sensitive index of bilingual balance (Dornic, 1979; Magiste, 1979; Potter et al., 1984).

As shown in Table 3, only correlations at the grade 9 level were significant. Two predominant patterns emerged from the correlation data. For the picture-word matching task, Verbal IQ scores of grade 9 extended French pupils were negatively correlated to French-English differential latency scores for the first eight trials (T1), the second eight trials (T2), and all trials. Although nonsignificant, a positive correlation between Verbal IQ and English latencies and a negative correlation between Verbal IQ and French latencies was found. This pattern of results suggests that the relationship between French-English differential latencies and Verbal IQ was a function of faster response speed in French (negative correlation) and slower response speed in English (positive correlation) with increasing

Verbal IQ. The other pattern occurred in the word-naming data. For grade 9 extended French pupils, there was a negative correlation between Verbal IQ and French-English differential latencies and between Verbal IQ and French latencies, which suggests that the decrement in French latencies was the main factor contributing to the reduced French-English differential latencies.

One other correlation was significant: Among grade 9 immersion pupils, Verbal IQ correlated negatively with English word-naming latencies. What this correlation may suggest is that Verbal IQ is related to word-naming skill in the language in which the student receives less of instruction. Thus, for extended French pupils, who receive most of their reading in English, students with higher Verbal IQ scores read words aloud faster in French than students with lower Verbal IQ scores. Conversely, for immersion pupils, who receive an extensive amount of practice reading French, pupils with higher Verbal IQ scores read English words faster. (Correlations for the category matching task are discussed later with other data from that task.)

Insert Table 3 about here

Table 3

Correlations Between Verbal IQ Scores and French, English, and French-English Differential Latencies for Picture-Naming, Word-Naming, and Picture-Word Matching Tasks

Group	n[a]	Latency Score: English	French	Differential
		Picture naming		
FI6	24	-.01	-.25	-.24
EF6	24	.15	-.18	.20
FI9	24	-.24	.15	.06
EF9	21	.32	-.21	-.31
		Word naming		
FI6	24	-.03	-.36	-.32
EF6	24	.08	-.25	-.30
FI9	24	-.42*	-.22	.19
EF9	21	.20	-.53*	-.71**
		Picture-word matching, first eight trials		
FI6	24	-.35	.04	-.37
EF6	24	-.10	-.25	-.23
FI9	24	-.16	.01	-.22
EF9	21	.36	-.36	-.54*
		Picture-word matching, second eight trials		
FI6	24	-.18	.02	.19
EF6	24	-.10	.06	.03
FI9	24	-.25	-.04	.39
EF9	21	.39	-.32	-.53**
		Picture-word matching, all trials		
FI6	24	-.26	.04	.34
EF6	24	-.15	-.08	.01
FI9	24	-.22	.01	.36
EF9	24	.38	-.30	-.51**

[a]Number of student for whom Verbal IQ scores were available.

*$p < .05$

**$p < .01$

B. Accuracy Data

Accuracy levels for the picture naming, word naming, and picture-word matching tasks with corresponding latencies are presented in Table 4. A ceiling effect was obtained in English on all tasks, invalidating further analyses. There was some variation among all four groups in French accuracy scores for the picture naming and category matching tasks, but for the word naming and picture-word matching tasks, the variation considerable only for the extended French pupils. Given this pattern, it was deemed inappropriate to pursue further discussion of the accuracy data in the results section, but relevant analyses are available in Appendix E. Effects due to possible trade-offs between accuracy and speed are discussed in the context of the latency data.

Insert Table 4 about here

C. Picture-Naming Latencies

For each subject, median latencies for naming three and six pictures were computed. Only error-free responses were used in the calculation of these median latencies. Two problems were encountered in the calculation and analysis of the slope and intercept data. First, latencies for naming six pictures in French were not available for all subjects. More extended French pupils were unable to name six pictures correctly than immersion pupils, which probably reflects the difference in French

Table 4

Mean Latencies and Accuracy in Percentages for the Picture-Naming (PN), Word-Naming (WN), and Picture-Word Matching (PWM) Tasks

Group	Task		
	PN	WN	PWM[a]
	English		
FI6			
Latency	2.50	1.44	5.63
Accuracy	99	99	99
EF6			
Latency	2.38	1.61	5.70
Accuracy	99	99	99
FI9			
Latency	2.40	1.60	5.23
Accuracy	99	100	99
EF9			
Latency	2.17	1.42	5.24
Accuracy	99	99	99
	French		
FI6			
Latency	3.99	2.03	5.90
Accuracy	91	100	99
EF6			
Latency	5.12	2.93	7.06
Accuracy	63	86	87
FI9			
Latency	2.92	1.69	5.21
Accuracy	96	100	99
EF9			
Latency	3.94	2.14	6.20
Accuracy	76	96	93

Note. Latency scores represent the time in seconds to name three words, three pictures, or match six picture-word pairs.

[a]All 16 trials of the picture-word matching task.

vocabulary between the two groups (see Table 5). Second, negative intercepts were obtained for all four groups in French and English. Due to these problems, no further analyses were performed on the slope and intercept data. The extreme number of negative intercepts was puzzling. One possible explanation is that the function for latencies is positively accelerated rather than linear. However, given that only two latency measures were collected, this hypothesis cannot be tested. One direction for future research would be to explore this hypothesis by including several lists of different length (e.g., 3, 6, 9, 12 and 15 words).

Insert Table 5 about here

A Grade (six, nine) by Program (French immersion, extended French), by Sex (male, female) by Testing Sequence (English-French, French-English) by Task Order (P1, W1, PW1) by Language (French, English) analysis of variance was computed with repeated measures on the last factor. As Task Order and Sex were nonsignificant and were not involved in any interactions, the analysis was repeated without these two variables. Sex was initially included in the analyses for all tasks. This variable has seldom been included in studies examining second language comprehension, production or memory organization (e.g., Caramazza & Brones, 1980; Magiste, 1982; Potter et al., 1984). In the present study, sex was included for exploratory purposes.

Table 5

Percentage of Subjects with Incomplete Latency or Negative Intercepts for the Pictue-Naming Task

Group	Incomplere Data for 6 Pictures	Negative Intercepts	Positive Intercepts	Intercepts M	Intercepts SD
FI6					
French	17	42	41	2.34	1.91
English	0	38	63	-.02	1.20
EF6					
French	63	16	21	.42	2.62
English	0	58	42	-.38	1.03
FI9					
French	0	83	17	.98	1.34
English	0	46	54	.23	.91
EF9					
French	38	50	12	-1.30	2.35
English	0	54	46	.05	.74

Main effects were found for Grade, $F(1,84) = 9.31$, $p < .01$, Program, $F(1,84) = 4.58$, $p < .05$, and Language, $F(1,84) = 69.54$, $p < .001$. Grade 6 students were slower than grade 9 students, extended French pupils were slower than immersion pupils, and pictures were named slower in French than English.

The three-way interaction between Grade, Program, and Language was nonsignificant, $F(1,84) < 1$, but the interactions were significant between Grade and Language, $F(1,84) = 9.88$, $p < .01$ and between Program and Language, $F(1,84) = 5.92$, $p < .05$. Tests of simple main effects were computed for these interactions. In the Grade by Language interaction, shown in Figure 3, French latencies decreased across grade level, $p < .001$, but English latencies did not, $F(1,168) = < 1$. In spite of the decrement in French, naming latencies in French were longer than in English at both grade levels, $ps < .001$.

Insert Figure 3 about here

In the Program by Language interaction, shown in Figure 4, longer French than English latencies were reported for immersion and extended French pupils, $p < .001$. However, immersion pupils named pictures in French faster than extended French pupils, $p < .001$. No differences attributable to program were found for English latencies, $F(1,168) < 1$.

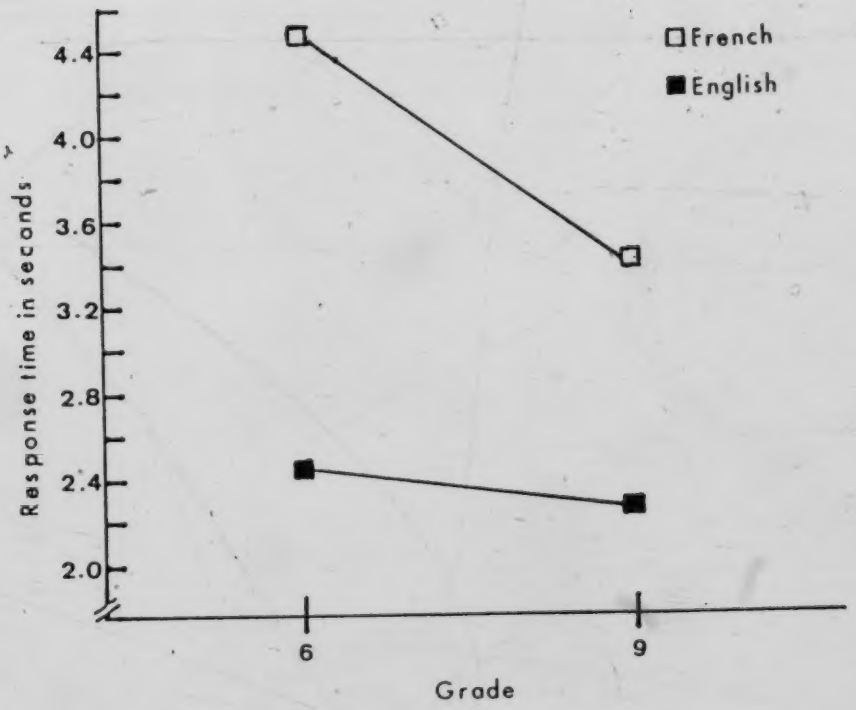


Figure 3. Mean response latencies for naming pictures as a function of Grade and Language.

Insert Figure 4 about here

French, English, and French-English differential latency means for each Program by Grade by Language condition are presented in Table 6. French-English differential latencies were examined for the four groups in spite of the nonsignificant three-way interaction because of a priori hypothesized relationships between duration of exposure (grade), intensity of exposure (program), and picture-naming latencies. Although grade 9 immersion pupils still took on the average 0.18 s longer to name pictures in French than English, the difference was nonsignificant, $F(1,84) = 1.71$. In contrast, grade 6 immersion, and grade 6 and 9 extended French pupils all took significantly longer to name pictures in French than in English, $ps < .001$ (see Figure 5).

Insert Table 6 about here

Insert Figure 5 about here

One of the most striking findings from Magiste's (1979, 1982) research was that first language latencies increased as exposure to the second language increased. A major purpose of the present study was to ascertain whether Magiste's findings could be replicated with immersion and extended French pupils.

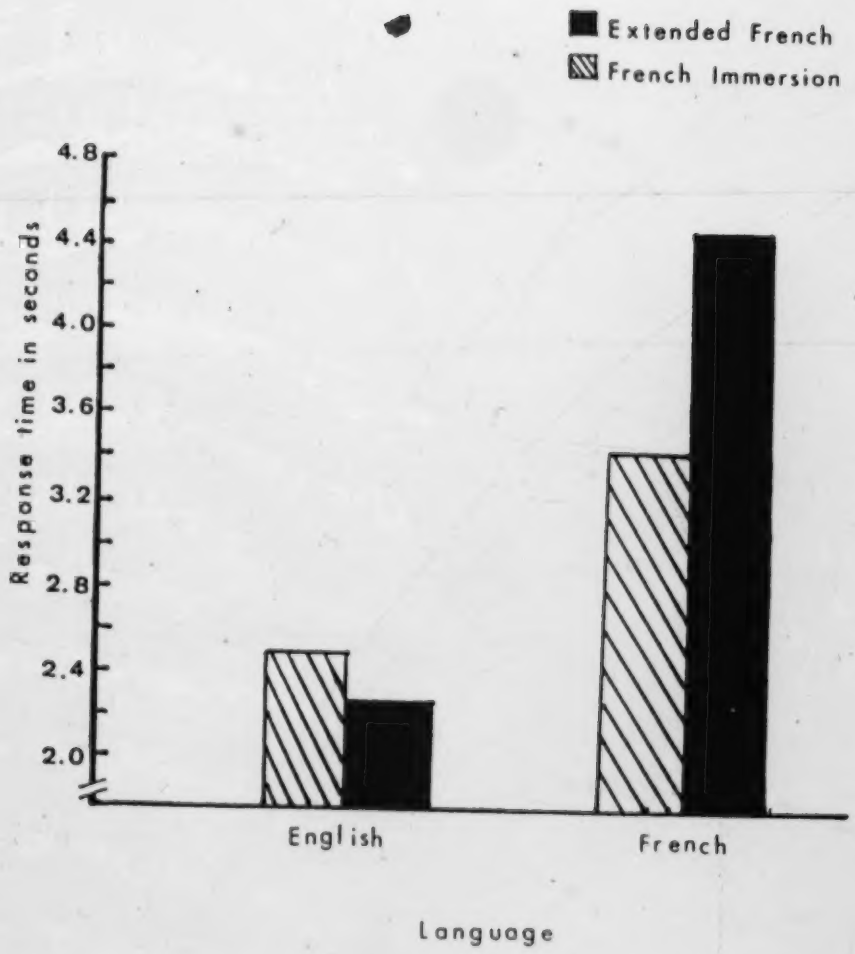


Figure 4. Mean response latencies for naming pictures as a function of Program and Language.

Table 6

French, English, and French-English Differential Latencies for Picture Naming, Word Naming, and Picture-Word Matching

Group	Language of Response		
	English	French	French-English
		Picture naming	
FI6	0.83	1.33	0.50[a]*
FI9	0.83	0.97	0.18
EF6	0.80	1.71	0.91*
EF9	0.72	1.31	0.59*
		Word Naming	
FI6	0.48	0.68	0.20*
FI9	0.53	0.56	0.03
EF6	0.54	0.98	0.44*
EF9	0.47	0.71	0.24*
		Picture-word matching, first eight trials	
FI6	0.96	1.00	0.04
FI9	0.88	0.88	0.00
EF6	0.96	1.21	0.25*
EF9	0.89	1.08	0.19*
		Picture-word matching, second eight trials	
FI6	0.94	0.98	0.05
FI9	0.87	0.87	0.00
EF6	0.95	1.18	0.23*
EF9	0.87	1.03	0.16*

Note. Latencies represent naming time for one word in seconds.

[a]Test of significant differences between English and French latencies.

*$p < .001$

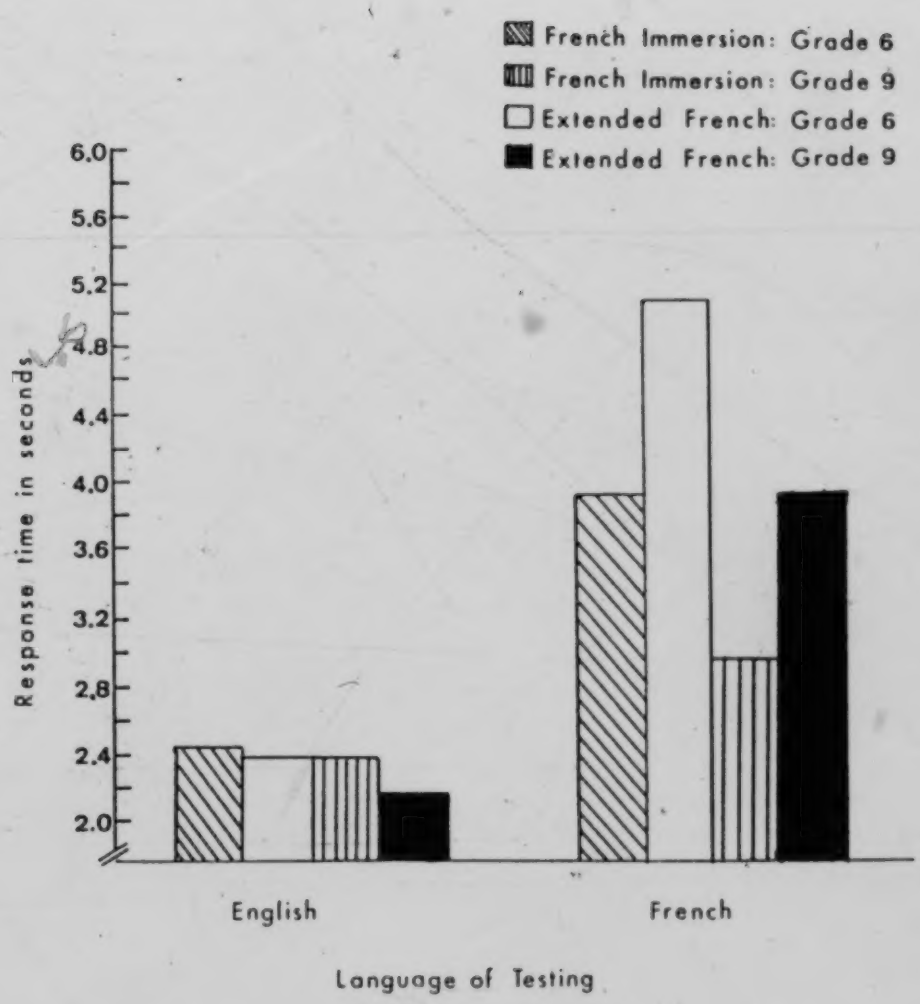


Figure 5. Mean response latencies for naming pictures as a function of Grade, Program and Language.

However, in the overall analysis of variance inclusion of the longer and more variable French latencies inflated the error term and reduced the sensitivity of the test. Thus, an analysis of variance was computed on English latencies only, and the a priori comparisons between duration of exposure (Grade) and intensity of exposure (Program) were examined using tests of simple effects. Also, a 95% confidence interval was constructed around the difference in means for each comparison.

The significance and range of possible true values for each of the four comparison are shown in the top portion of Table 7. For all four comparisons, negative differences imply that the group with less exposure (grade 6 or extended French) had longer latencies, and positive differences imply that the group with more exposure (grade 9 or immersion) had longer latencies. The first two comparisons were used to assess the effects of duration of exposure for students in each program. In picture naming, the negative mean difference values for the first two comparisons indicate a decrease in latencies across grades for immersion and extended French pupils, which is opposite to Magiste's predictions. The decrement across grade, however, was not significant for extended French, $F(1,92) = 2.22$, or for immersion pupils, $F(1,92) < 1$. The third and fourth comparisons assess effects of intensity of exposure at each grade. Positive values indicate that students with more intense exposure to French (immersion) have longer response latencies than students with minimal exposure (extended French). Thus, positive mean difference values support Magiste's previous findings, whereas

negative values contradict them. At both grade levels, the mean difference values were in the direction predicted by Magiste. The difference between the groups was not significant at grade 6, $F(1,92) < 1$, or at grade 9; $F(1,92) = 2.46$.

Insert Table 7 about here

Discussion. French picture-naming latencies decreased as the duration or number of years of exposure to French increased, regardless of whether students were enrolled in an intensive French program (immersion) or less intense program (extended French). This decrease in French latencies across grade is confounded partially by decrements usually found to correspond with age. However, if age was the primary factor responsible for the reduction in French decoding latencies, then a significant decrement would have been expected for English latencies, also. Instead, only the decrease in French latencies across grade was significant, which suggests that duration of exposure was an important factor.

Contrary to Magiste's (1979, 1982) findings, English latencies did not increase as years of exposure increased for students in either program. Rather, a decrease across grade was found for extended French and immersion pupils, although the decrease was larger for extended French pupils. As previously stated, if the decrement across grade differs for the two groups, this suggests the possibility that age effects may have been

Table 7

Mean English Latency Differences and Confidence Intervals for Picture Naming, Word Naming, and Picture-Word Matching

Comparison	Mean Difference	95% Confidence Interval	
Picture naming			
Grade 9 - 6, Extended French[a]	-.22	-.68	.23
Grade 9 - 6, Immersion	-.10	-.55	.35
Immersion - Extended, Grade 6[b]	.11	-.34	.57
Immersion - Extended, Grade 9	.23	-.22	.69
Word naming			
Grade 9 - 6, Extended French	-.19**	-.37	-.01
Grade 9 - 6, Immersion	.16*	-.02	.34
Immersion - Extended, Grade 6	-.17*	-.36	.01
Immersion - Extended, Grade 9	.17*	-.01	.36
Picture-word matching			
Grade 9 - 6, Extended French	-.49**	-.94	-.04
Grade 9 - 6, Immersion	-.41*	-.86	.03
Immersion - Extended, Grade 6	-.07	-.52	.37
Immersion - Extended, Grade 9	.02	-.43	.46

Note. Mean latency differences are reported in seconds.

[a]Latencies for grade 6 pupils are subtracted from grade 9 pupils. Negative differences imply that grade 6 pupils were slower. [b]Latencies for extended French pupils are subtracted from immersion pupils and negative differences imply that extended French pupils were slower.

*$p < .10$

**$p < .05$

modified by differences between the two groups with respect to duration and intensity of exposure. Specifically, it is possible that the smaller decrement across grade for immersion students reflects a decrease associated with age and an increase associated with exposure, whereas the larger decrement for extended French pupils reflects primarily a decrement associated with age. However, the difference between the size of the decrement across grade for immersion and extended French pupils was small. Thus, the only conclusion that can be drawn is that the Grade effects do not support Magiste's duration effects hypothesis, but neither do they provided a very strong refutation.

Magiste (1979, 1982) also reported that bilinguals took longer to name pictures in their first language than nonbalanced bilinguals or monolinguals. In the present study, the differences in latencies between immersion and extended French pupils were nonsignificant but in the direction predicted by Magiste's data.

D. Word-Naming Latencies

Median latencies for three and six words were computed for each subject. As indicated in Table 8, there were a sufficient number of three- and six-word latencies to calculate slopes, but the high percentage of negative intercepts made interpretation of the slope data questionable. All further analyses were computed on median latencies for naming three words.

Insert Table 8 about here

Given that there were several significant correlations with Verbal IQ (Table 3), a Grade (six, nine) by Program (French immersion, extended French) by Sex (male, female) by Testing Sequence (English-French, French-English) by Task Order (P1, W1, PW1) by Language (French, English) analysis of covariance was computed with repeated measures on the last factor. Verbal IQ was the covariate. The assumption of homogeneity of regression coefficients was satisfied.

As with picture-naming, three main effects were identified: Grade, $F(1,45) = 4.7$, $p < .05$, Program, $F(1,45) = 23.31$, $p < .001$, and Language, $F(1,45) = 121.23$, $p < .001$. Grade 6 students were slower than grade 9 students, extended French pupils were slower than immersion students, and words were named more slowly in French than in English.

The three-way interaction between Grade, Program, and Language was nonsignificant, $F(1,45) < 1$, but interactions were significant between Grade and Language, $F(1,45) = 22.41$, $p < .001$, Program and Language, $F(1,45) = 29.72$, $p < .001$, and Program and Grade, $F(1,45) = 4.35$, $p < .05$. Tests of simple effects were computed to explore these interactions. With respect to the Grade by Language interaction, there was a significant decrement across grade in French, $p < .001$, but not in English, $F(1,90) < 1$. At both grade levels English words were read aloud faster than French words, $ps < .001$, as shown

Table 8

Percentage of Subjects with Incomplete Data or Negative Intercepts for the Word-Naming Task

Group	Incomplete Data for 6 Words	Negative Intercepts	Positive Intercepts	Intercepts M	Intercepts SD
FI6					
French	0	25	75	.47	.35
English	0	42	58	.31	.22
EF6					
French	4	46	50	-.04	1.08
English	0	17	83	.28	.48
FI9					
French	0	83	17	.14	.33
English	0	38	62	.44	.22
EF9					
French	0	46	54	.09	.57
English	0	50	50	.33	.23

in Figure 6.

Insert Figure 6 about here

Examination of the Program by Language interaction revealed that students in the French immersion program read French words faster than extended French pupils, $p < .001$, but no difference attributable to Program was found for English latencies, $F(1,90) < 1$. However, as shown in Figure 7, pupils in both programs read words faster in English than French, $ps < .01$.

Insert Figure 7 about here

The Grade by Program interaction, shown in Figure 8, can best be understood in light of the two previous interactions. There was a significant decrease in word-naming latencies across Grade for extended French pupils, $p < .001$, but not for immersion students, $F(1,45) < 1$, which might possibly reflect a floor effect in French for immersion pupils. At grade 6, immersion pupils were faster than extended French pupils, $p < .001$, but the difference was nonsignficant at grade 9, $F(1,45) = 1.63$.

Insert Figure 8 about here

A cross-language facilitation effect was evident upon examination of the simple main effects involved in the Grade by

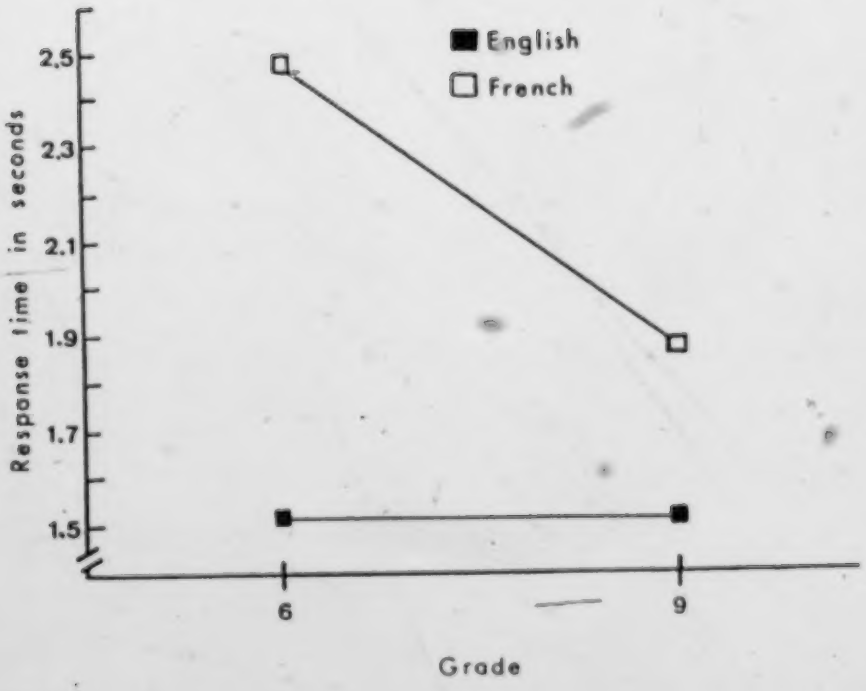


Figure 6. Mean response latencies for reading words aloud as a function of Grade and Language

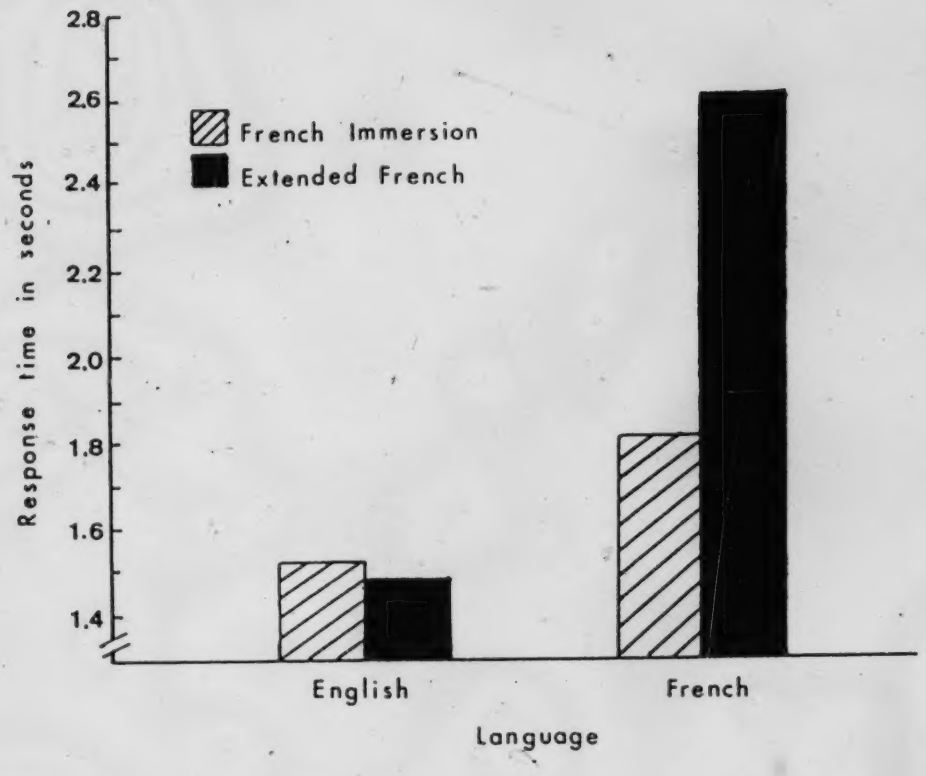


Figure 7. Mean response latencies for reading words aloud as a function of Program and Language

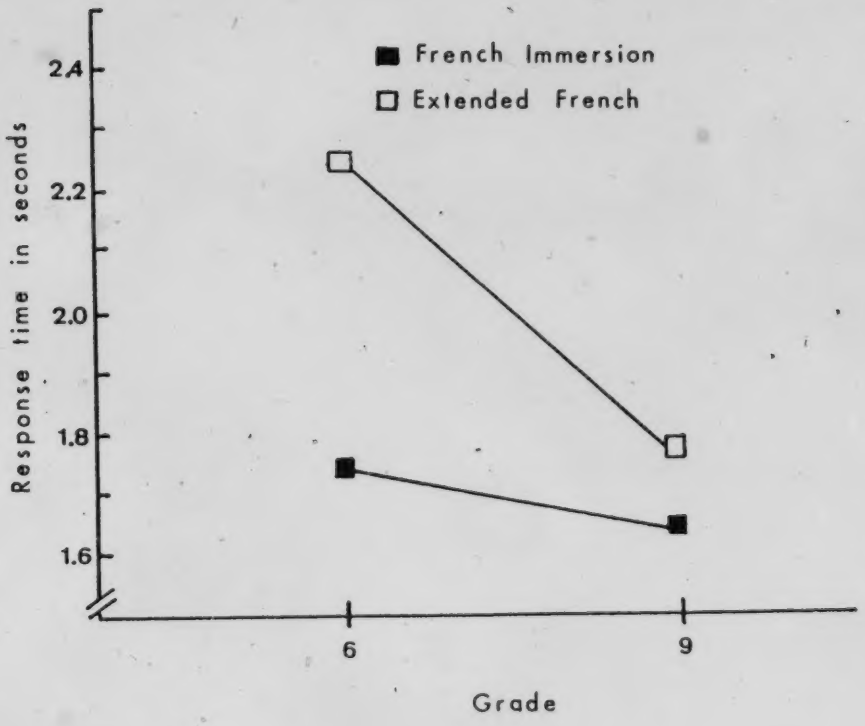


Figure 8. Mean response latencies for reading words aloud as a function of Grade and Program.

Testing Sequence by Language interaction, $F(1,45) = 8.43$, $p < .01$. As shown in Figure 9, there was no effect of Test Sequence on English latencies at the grade 6, $F(1,45) = 1.67$, or grade 9 level, $F(1,45) < 1$. However, in French a facilitory effect of Sequence was found at grade 9, $p < .01$, such that response speed was faster when words were presented first in English. Although Ehri and Ryan (1980) used pictures, they reported a similar trend. Presentation of pictures first in the dominant language facilitated performance in the weaker language, but the converse was not found.

Insert Figure 9 about here

Program, Sex, and Task Order interacted, $F(2,45) = 2.71$, $p < .05$. As shown in Figure 10, with the exception of female extended French pupils, no significant difference was found between Task Order W1 and PW1. This outcome suggests that receiving the comprehension task first did not facilitate performance. Among female extended French pupils, there was a significant facilitory effect of Task Order PW1 when compared to W1, $p < .05$, but this finding is questionable because the high mean latency value for W1 resulted from the high score of one of the four subjects in that cell. Comparison of Task Order P1 to W1 revealed that although all but female extended French pupils evidenced longer latencies when picture naming preceded word naming, the difference was significant only for female immersion pupils, $p < .01$, and approached significance for male immersion

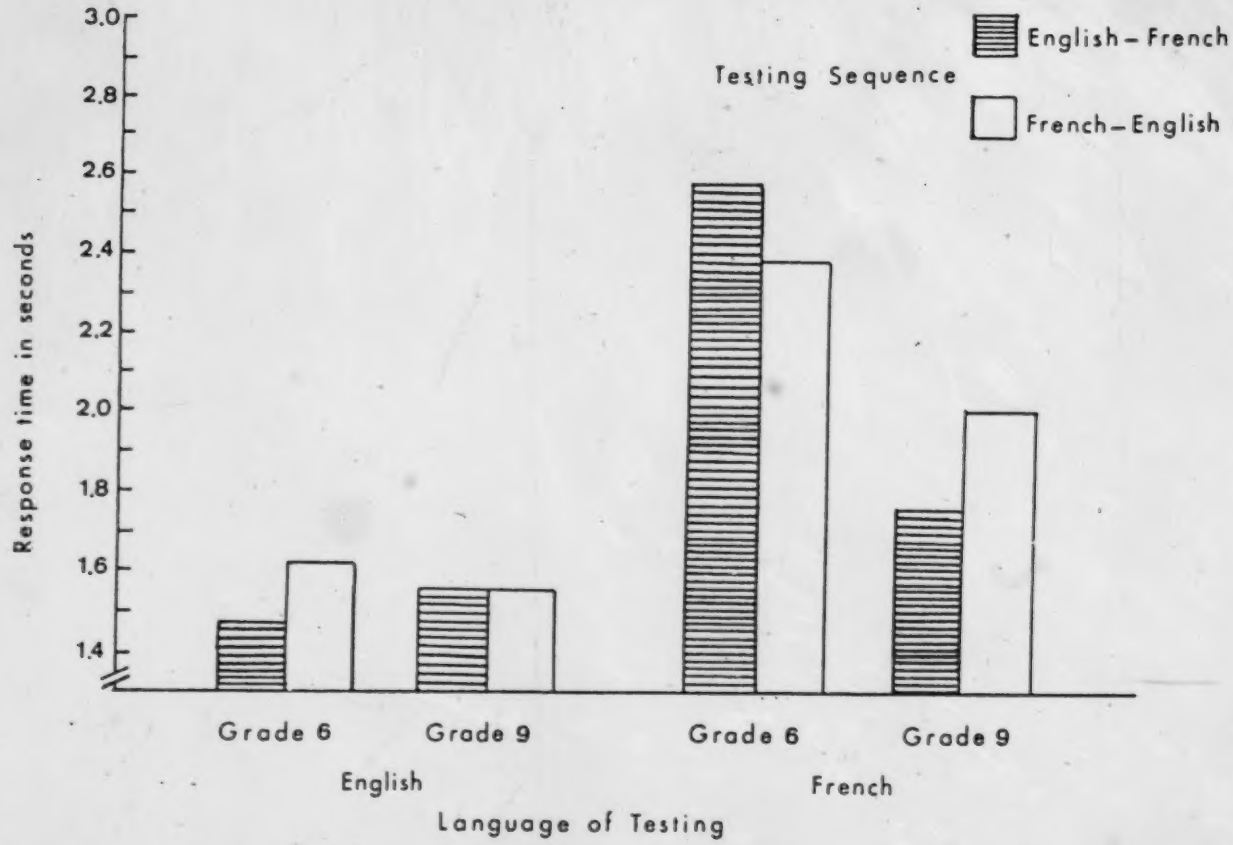


Figure 9. Mean response latencies for reading words aloud as a function of Grade, Language, and Testing Sequence.

pupils, $F(1,45) = 3.97$, $p < .10$. In summary, receiving picture naming prior to word naming actually increased response latencies. Receiving the comprehension task prior to the word naming task, had either little or no facilitory effect.

Insert Figure 10 about here

As was the case for the picture naming data, French-English differential latencies were examined for each of the four groups in spite of the nonsignficant three-way interaction (Table 6). Grade 9 immersion pupils appeared to be reaching a point of bilingual balance, as the mean French-English differential latency value was only 0.03 s, $F(1,45) = .59$. In contrast, grade 6 immersion, grade 6 extended French, and grade 9 extended French pupils named English words faster than French, $ps < .001$.

A separate analysis of covariance was computed on English latencies and the results are shown in the middle portion of Table 7. The rationale and procedures were the same as those outlined for picture naming. The Grade by Program interaction was significant, $F\ (1,88) = 6.64$, $p < .05$. As indicated in the first two comparisons, the effects of duration of exposure differed for each group. The positive mean difference value for immersion pupils, although only marginally significant, $F(1,88) = 3.28$, $p < .10$, is in the direction predicted by Magiste. The significant negative difference for extended French pupils, $F(1,88) = 4.36$, $p < .05$, is consistent with decrements associated with age but not with Magiste's

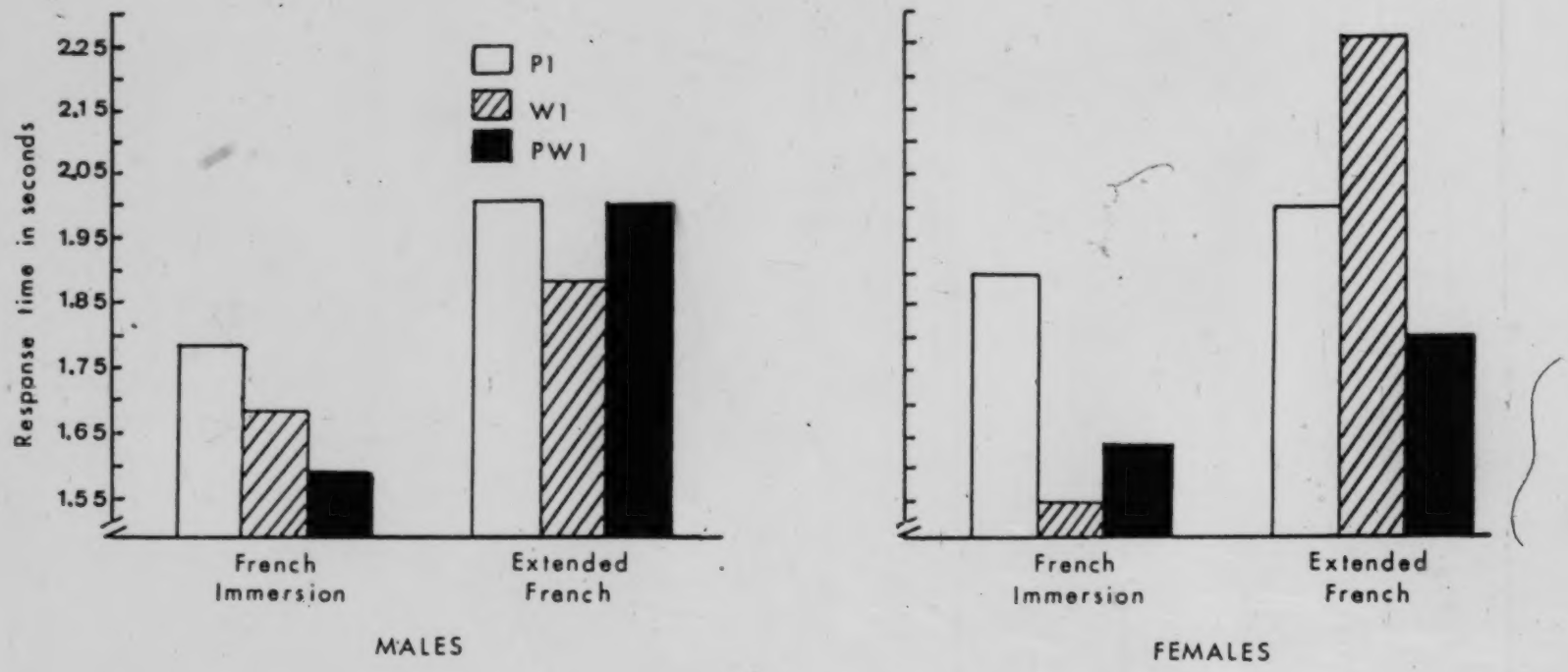


Figure 10. Mean response latencies for reading words aloud as a function of Program, Sex and Task Order.

predictions. The third and fourth comparisons test the effects of intensity of exposure. At grade 9, immersion pupils were marginally slower than extended French pupils, $F(1,88) = 3.79$, $p < .10$. This positive mean difference value for grade 9 pupils is in the direction predicted by Magiste. In contrast at grade 6, the mean difference was negative and marginally significant, $F(1,88) = 3.79$, $p < .10$.

Discussion. The decrement in French latencies associated with increased duration of exposure was larger for extended French than for immersion pupils. However, immersion students named French words faster than same-aged extended French pupils at the grade 6 and 9 level.

Among immersion pupils, there was a marginally significant increase in English latencies from grades 6 to 9, but among extended French pupils there was a decrease across grade level. Thus, an increase in English response speed associated with duration of exposure was found, but only for pupils who had received intense exposure to French. These results only partially support Magiste's (1979) predictions.

Grade 9 immersion pupils, who had reached a point of bilingual balance, took longer to read words in English than grade 9 extended French pupils but the difference was only marginally significant. The direction of this effect is consistent with Magiste's (1979, 1982) data, which indicates that balanced bilinguals take longer to read words in their first language than nonbalanced bilinguals or monolinguals. At grade 6, however, immersion pupils were faster than extended French

pupils. This finding was unexpected and is not consistent with previous research findings (Magiste, 1979, 1982).

E. Picture-Word Differential Latencies

Inspection of Table 9 reveals a range in English picture-word differential scores that is relatively consistent with the 0.2 s to 0.3 s difference reported for monolinguals (e.g., Fraisse, 1960; Huttenlocher & Kubicek, 1985; Snodgrass, 1984) and bilinguals (Magiste, 1979; Potter et al., 1984). However, in the second language, the range was from 0.41 s (grade 9 immersion) to 0.7 s (grade 6 extended French).

Insert Table 9 about here

To explore the differences between picture-word latencies, a Grade (six, nine) by Program (French immersion, extended French) by Language (French, English) analysis of variance was computed on picture-word differential latencies with repeated measures on the last factor. Only the main effect of Language was significant, $F(1,92) = 29.6$, $p < .001$. The picture-word differential was significantly larger in French than English.

F. Picture-Word Matching Latencies

Median latencies for sorting six picture-word pairs were calculated for each subject. To assess practice effects, median response time was computed for the first 8 trials (T1), second 8 trials (T2), and all trials.

Table 9

Picture Naming, Word Naming, and Differential Picture-Word Latencies

Group	Latencies		
	Picture	Word	Picture-word differential
		English	
FI6	0.83	0.48	0.35
EF6	0.90	0.54	0.26
FI9	0.90	0.53	0.26
EF9	0.72	0.47	0.25
		French	
FI6	1.33	0.68	0.66
EF6	1.71	0.98	0.73
FI9	0.97	0.56	0.41
EF9	1.31	0.71	0.60

Note. Latencies represent naming time for one word or picture in seconds.

Given that there were several significant correlations between Verbal IQ and differential latency scores, a Grade (six, nine) by Program (French immersion, extended French) by Sex (male, female) by Testing Sequence (English-French, French-English) by Task Order (P1, W1, PW1) by Trial (T1, T2) by Language (French, English) analysis of covariance was computed with repeated measures on the last two factors. The assumption of homogeneity of regression coefficients was satisfied.

Main effects were found for Grade, $F(1,45) = 4.47$, $p < .05$, Program, $F(1,45) = 10.84$, $p < .01$, Sex, $F(1,45) = 8.27$, $p < .01$, Test Sequence, $F(1,45) = 4.82$, $p < 05$, Trial, $F(1,45) = 21.76$, $p < .001$, and Language $F(1,45) = 61.71$, $p < .001$. Grade 6 students were slower than grade 9 pupils, extended French pupils were slower than immersion students, males were slower than females, latencies for Test Sequence French-English were longer than for English-French, latencies for the first eight trials were longer than for the last eight trials, and French latencies were longer than English latencies.

Program interacted with Trial, $F(1,45) = 5.97$, $p < .05$. Examination of the simple main effects revealed a practice effect (i.e., reduced latencies from T1 to T2) for extended French pupils, $p < .001$, but not for immersion students, $F(1,45) = .001$ (see Figure 11).

Insert Figure 11 about here

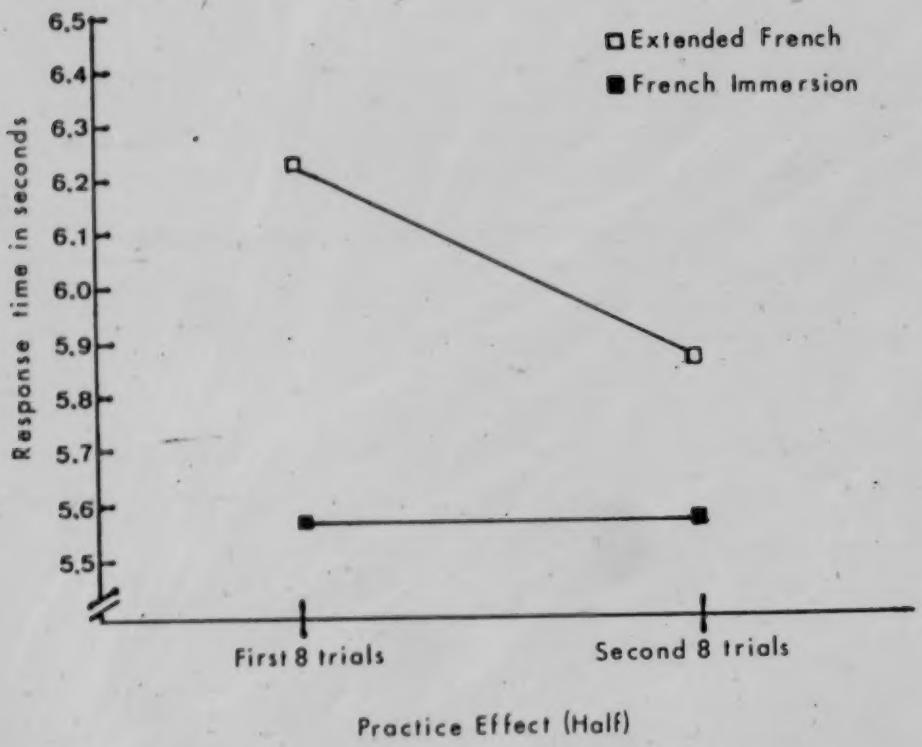


Figure 11. Mean response latencies for matching picture–word pairs as a function of Program and Practice across trials.

Language and Testing Sequence interacted, $F(1,45) = 66.14$, $p < .001$. As shown in Figure 12, French latencies were faster when the task was administered first in English (Testing Sequence English-French), $p < .001$. Although a similar decrement was noted in English latencies following administration of the tasks first in French (Testing Sequence French-English), it was nonsignificant, $F(1,90) = 3.93$. This cross-language facilitation effect from the stronger to weaker language was also found in the word naming data for grade 9 pupils and in prior research (Ehri & Ryan, 1980).

Insert Figure 12 about here

Program interacted with Language, $F(1,45) = 45.43$, $p < .001$. As shown in Figure 13, the difference between English and French latencies was significant for extended French pupils, $p < .001$, but not for immersion pupils, $F(1,45) < 1$. Immersion students comprehended French word meanings faster than extended French pupils, $p < .001$, but the difference between the groups in English latencies was marginal, $F(1,90) < 1$.

Insert Figure 13 about here

Grade, Program, Sex, Task Order, Testing Sequence, and Language were involved in an interaction, $F(2,45) = 4.46$, $p < .05$. Due to the complexities of interpreting a six-way interaction, the main purpose was to determine whether any of the

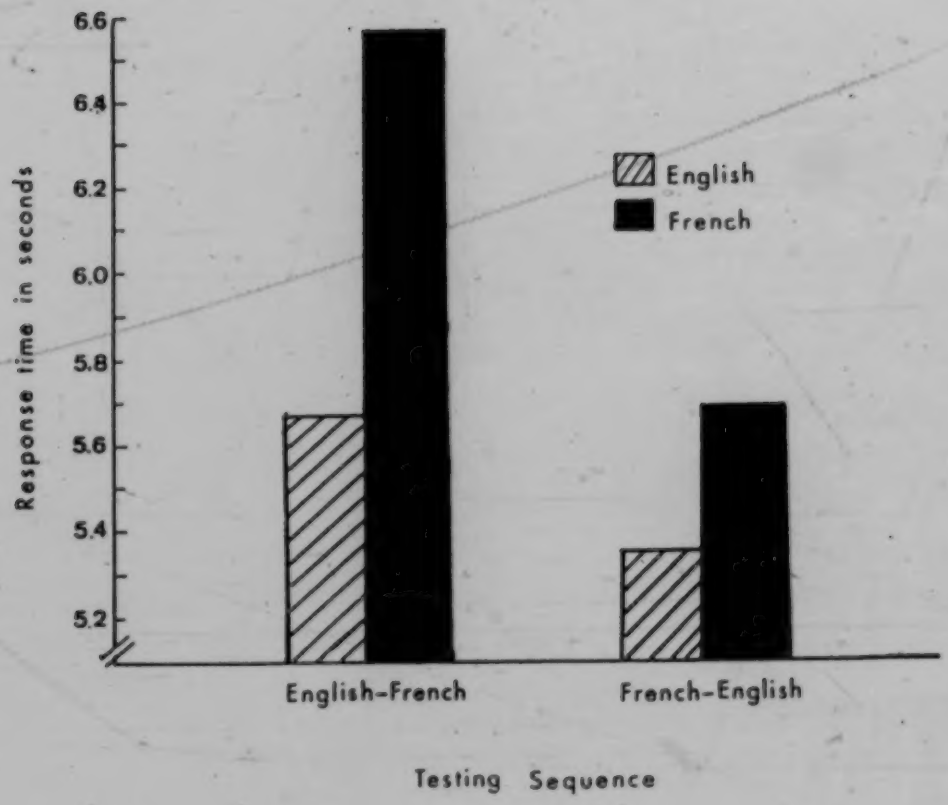


Figure 12. Mean response latencies for matching picture-word pairs as a function of Testing Sequence and Language.

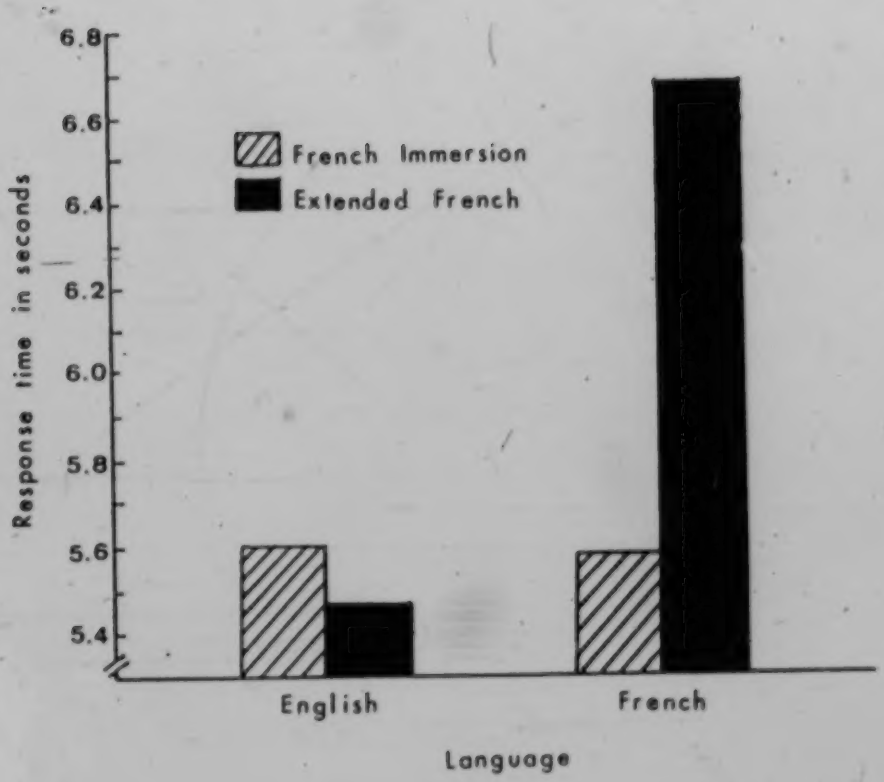


Figure 13. Mean response latencies for matching picture-word pairs as a function of Program and Language

significant two-way interactions changed at some level of the other variables. The finding of interest was that for males and females in all four groups, there was a consistent minimal decrease in English latencies when the picture-word pairs were presented first in French (sequence French-English), but the difference was not significant. In contrast, French latencies decreased significantly when the picture-word pairs were presented first in English, for all but grade 9 male immersion students, $F(1,90) < 1$.

Examination of the French-English differential encoding latencies for each of the four groups (Table 6) revealed that both grade 6 and 9 immersion pupils, with differential latencies of 0.05 s and -0.004 s respectively, appeared to have reached a point of bilingual balance on this task. In contrast, for grade 6 and 9 extended French pupils a significant French-English differential was found, $ps < .01$, indicating both groups were still English-dominant. Across grade, the decrement in French latencies was significant for extended French and immersion pupils, $ps < .01$. At grade 6, immersion pupils matched picture-word pairs faster than same-aged students in the extended French program, $p < .001$. However, a comparison of the two instructional groups at grade 9 revealed no significant differences, $F(1,45) < 1$.

A separate analysis of covariance was computed for English picture-word matching latencies which are presented in the bottom portion of Table 7. As indicated in the first two comparisons, there was a decrease across grade for immersion, $F(1,88) = 3.24$,

$p < .10$, and extended French pupils, $F(1,88) = 4.52$, $p < .05$. These results are not consistent with Magiste's previous findings. Effects for intensity of exposure, tested in the third and fourth comparison, were nonsignificant for immersion, $F(1,88) < 1$, and extended French pupils, $F(1,88) < 1$.

Discussion. The results from the comprehension task paralleled those from picture naming. French latencies decreased across grade level for immersion and extended French pupils. In word naming, a decrement in French latencies across grade was found only for extended French students. Comparison of French picture-word matching latencies across instructional groups revealed that immersion pupils were faster than extended French pupils at grade 6, but not at grade 9, which may reflect a floor effect for immersion pupils. A similar pattern of results was found in the word-naming data, but for picture naming immersion students were faster than extended French pupils at both grade levels.

The predicted increase in first language latencies associated with increased duration of exposure (Magiste, 1979, 1982) was not found for immersion pupils. Rather, a decrement across grade of similar magnitude was reported for immersion and extended French pupils.

The second prediction generated from Magiste's (1979, 1982) work, that nonbalanced bilinguals would have faster first language latencies than balanced bilinguals, was not supported. The difference in English latencies between immersion and extended French pupils was not significant at either grade.

Given that grade 6 and 9 immersion pupils had reached a point of bilingual balance, the present data do not support Magiste's earlier finding. These results, however, support Favreau and Segalowitz's (1982) finding that subjects can demonstrate equivalent comprehension scores in each language, yet still read faster in their first language.

G. Category Matching Latencies

For each subject, data from error-free trials were used to compute median latencies in milliseconds for match and mismatch trials. Response latencies for the first (H1) and second half (H2) of the trials were computed to allow for examination of practice effects. Median latencies and percentage accuracy scores for each group are shown in Table 10. Except where noted, conclusions from latencies were not confounded by a speed-accuracy trade-off. As shown in Table 11, none of the correlations between Verbal IQ and Language Condition (English-English, French-French, English-French, and French-English) latencies were significant.

Insert Table 10 about here

Insert Table 11 about here

Table 10

Median Latencies and Accuracy Scores for the Category Matching Task

Group	Language Condition			
	EE	FF	EF	FE
	Matched trials			
FI6				
Latency[a]	1.41	1.62	1.55	1.56
Accuracy[b]	93	90	96	92
EF6				
Latency	1.69	2.23	1.98	2.04
Accuracy	97	86	87	84
FI9				
Latency	1.24	1.38	1.33	1.34
Accuracy	96	94	97	91
EF9				
Latency	1.16	1.56	1.34	1.41
Accuracy	91	88	91	86
	Mismatched trials			
FI6				
Latency	1.49	1.73	1.61	1.66
Accuracy	97	96	97	96
EF6				
Latency	1.85	2.46	2.06	2.10
Accuracy	97	81	89	85
FI9				
Latency	1.31	1.49	1.40	1.46
Accuracy	98	93	97	92
EF9				
Latency	1.27	1.49	1.39	1.43
Accuracy	94	85	89	92

Note. EE = English-English word pairs; FF = French-French word pairs; EF = English-French word pairs; FE = French-English word pairs.

[a]Latencies in seconds. [b]Accuracy in percentages

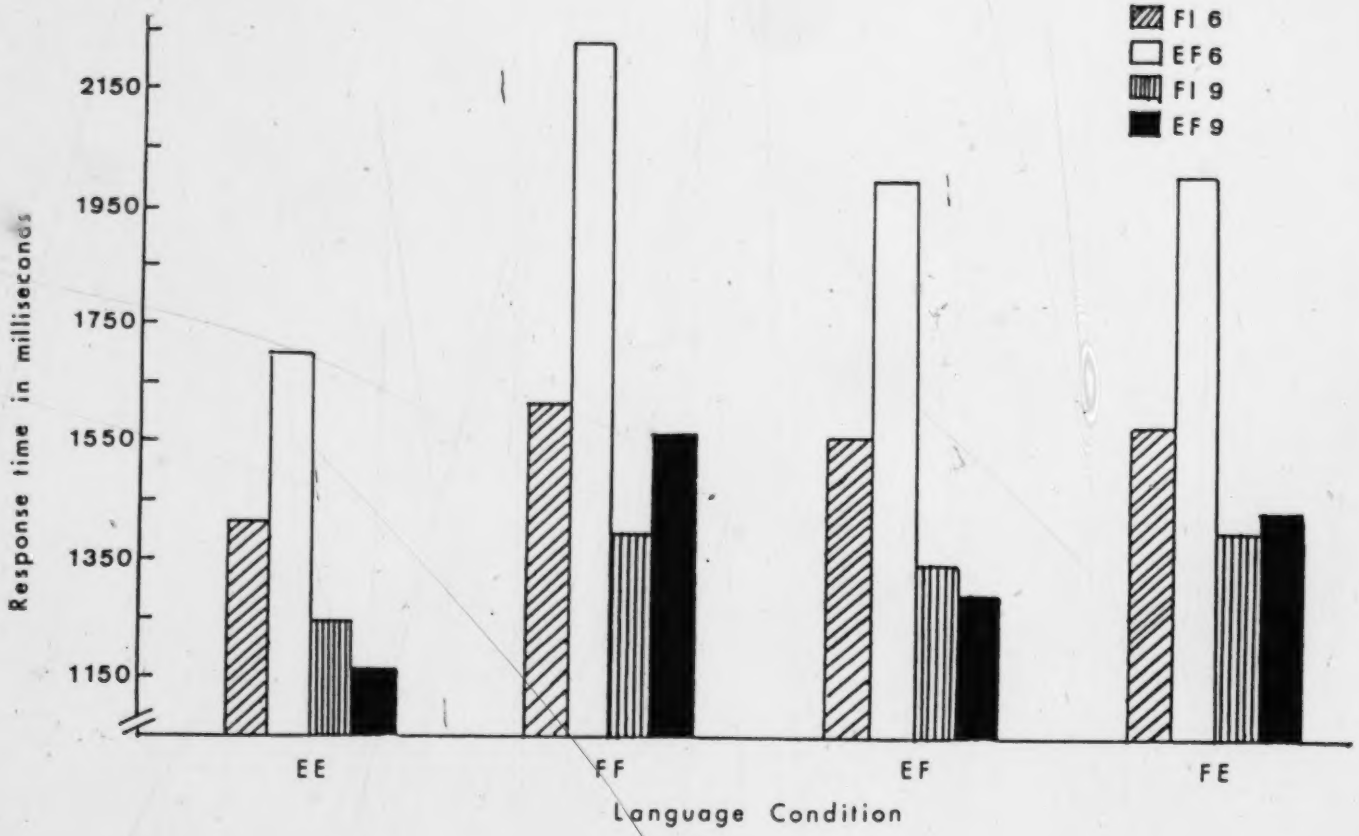


Figure 14. Mean response latencies for all groups at each language condition for matched trials on the Category Matching task.

A Grade (six, nine) by Program (French immersion, extended French) by Language Condition (English-English, French-French, English-French, French-English) by Half (H1, H2) by Match (match, mismatch) analysis of variance was computed with repeated measures on the last three factors. There was a main effect for Match, $F(1,80) = 30.21$, $p < .001$ and several interactions involving Match, including a Grade by Program by Match interaction, $F(1,80) = 4.35$, $p < .05$, and a Grade by Program by Match by Half interaction, $F(1,80) = 5.01$, $p < .05$. To, facilitate interpretation, separate analyses were computed for matched and mismatched latencies.

Memory models for matched latencies. The two main goals of this experiment were (a) to determine which of the four proposed models of bilingual semantic memory best fit the data for each of the four groups and (b) to assess whether bilingual semantic memory organization undergoes structural change as competency in the second language increases.

For each of the four groups, a Sex (male, female) by Half (H1, H2) by Language Condition (English-English, French-French, English-French, French-English) analysis of variance was computed with repeated measures on the last two factors. A main effect for Language Condition was found for grade 6 immersion pupils, $F(3,60) = 5.02$, $p < .05$, for grade 6 extended French pupils, $F(3,60) = 17.2$, $p < .001$, for grade 9 immersion pupils, $F(3,60) = 4.28$, $p < .05$, and for grade 9 extended French pupils, $F(3,60) = 21.16$, $p < .001$. The pattern of scores across Language Condition for all groups in shown in Figure 14.

Insert Figure 14 about here

Main effects of Language Condition were analyzed using the three orthogonal multiple comparisons designed to test the predictions of each of the four models of bilingual memory (Table 1). Examination of these comparisons, which are summarized in Table 12, reveals that: (a) latencies were longer in the French-French than English-English condition for all groups, ps < .01; (b) differences between the same (English-English, French-French) and different (English-French, French-English) language conditions did not reach significance for any of the four groups; and (c) differences between the two different language conditions (English-French, French-English) were nonsignificant for all four groups, indicating no asymmetry effects in the direction of translation. This pattern of results is consistent with a One-Store Different Encoding model of bilingual semantic memory.

Insert Table 12 about here

Language Condition interacted with Half for grade 9 immersion pupils, $F(3,60) = 4.28$, $p < .05$. The effects of Language Condition at H1 mirrored the pattern of results previously reported for all groups. As shown in Figure 15, at H2 there was no significant difference between English-English and French-French latencies ($M = 0.04$ s), $F(1,240) < 1$, or between

Table 11

Correlations Between Verbal IQ and Category Matching Latencies

Group	n	Language Condition			
		EE	FF	EF	FE
		Matched trials			
FI6					
H1	24	-.26	-.26	-.28	-.19
H2	24	-.17	-.01	-.15	-.16
HA	24	-.23	-.02	-.19	-.17
EF6					
H1	24	.26	-.26	-.28	-.19
H2	24	.14	-.01	-.15	-.16
HA	24	.10	-.02	-.19	-.17
FI9					
H1	24	.01	-.38	.08	.06
H2	24	.02	.03	.12	.14
HA	24	.04	-.02	.15	.16
EF9					
H1	21	.04	-.06	.05	.07
H2	21	.26	.24	-.33	.13
HA	21	.15	.05	-.02	.12
		Mismatched trials			
FI6					
H1	24	-.29	-.29	-.38	-.29
H2	24	-.34	-.26	-.17	-.12
HA	24	-.33	-.24	-.24	-.17
EF6					
H1	24	-.12	-.29	-.38	-.29
H2	24	.01	-.26	-.17	-.12
HA	24	-.01	-.24	-.24	-.17
FI9					
H1	24	-.13	-.17	-.05	-.28
H2	24	.10	-.03	.09	-.04
HA	24	.06	-.17	.04	-.15
EF9					
H1	21	.11	-.05	.03	.05
H2	21	.26	.17	.25	.11
HA	21	.16	.01	.11	.09

Note. H1 = first half of trials; H2 = second half of trials; HA = all trials.

Table 12

Observed Values for Prediction Equations for Models of Bilingual Semantic Memory

Comparison	Group			
	FI6	FI9	EF6	EF9
	Match trials			
(a) FF-EE	0.21**	0.14*	0.54**	0.40**
(b) $\frac{FE+EF}{2} - \frac{FF+EE}{2}$	0.09	0.10	0.11	0.03
(c) FE-EF	0.01	0.05	0.06	0.07
	Mismatch trials			
(a) FF-EE	0.24**	0.19**	0.61**	0.22**
(b) $\frac{FE+EF}{2} - \frac{FF+EE}{2}$	0.05	0.06	-0.06	0.03
(c) FE-EF	0.05	0.07	-0.06	0.08

Note. All reported values are in seconds.

*$p < .01$

**$p < .001$

English-French and French-English latencies (M = 0.08 s), $F(1,240) = 2.7$. However, the difference between same and different language trials was significant, $p < .001$, which suggests the presence of a translation parameter and thus is consistent with a Two-Store Equal Encoding model.

Insert Figure 15 about here

The results for the second half of the trials for grade 9 immersion pupils were unexpected. It is implausible to presume that, half way through a task subjects switched from one to two permanent stores. One possible explanation is that grade 9 immersion students recognized quickly that only a few words in each language were being repeated. Consequently, they may have developed a temporary, task-specific store consisting of the seven French words. As shown in Table 13, the main change in latencies on the second half of the trials was a reduction in French latencies, which could have resulted from setting up a temporary store in French. To verify this explanation it would be useful to have subjects (a) complete a category matching task with the same number of trials but more word pairs, and (b) do the same task with more trials. If subjects were found to use a one-store model in the former case, and a two-store model in the latter case, then it could be argued that bilinguals can set up a temporary store when it is to their advantage on a particular task.

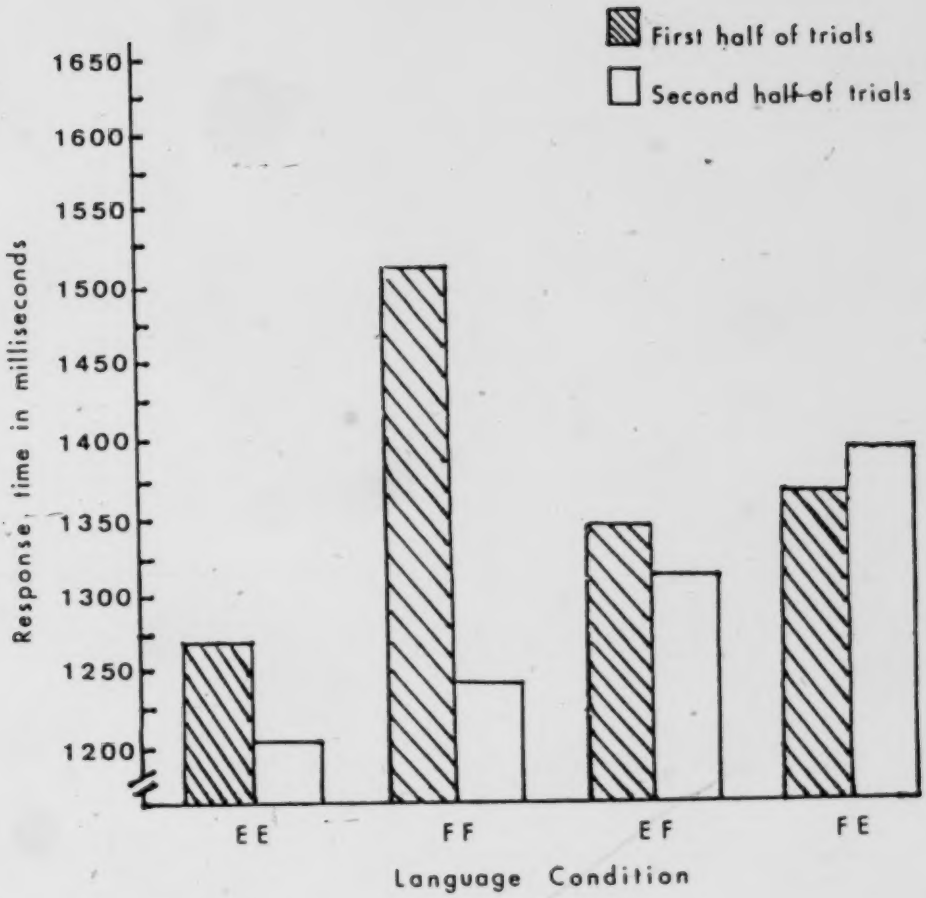


Figure 15. Mean response latencies for French Immersion Grade 9 pupils on the matched trials of the Category Matching task as a function of Language Condition and Trial.

Insert Table 13 about here

Mismatched latencies. The procedure for data analysis followed that outlined for matched latencies. The only significant finding was a main effect of Language Condition for grade 6 immersion students, $F(3,60) = 8.7$, $p < .01$, grade 6 extended French pupils, $F(3,60) = 15.66$, $p < .001$, grade 9 immersion pupils, $F(3,60) = 8.78$, $p < .01$, and grade 9 extended French pupils, $F(3,60) = 9.46$, $p < .01$. The results of the orthogonal comparisons testing the predictions of the four memory models are presented in Table 12. The pattern of results was similar for all four groups: (a) latencies were longer in the French-French than English-English conditions, $ps < .01$; (b) differences between the same and different language conditions were nonsignificant; and (c) differences between the two different language conditions were nonsignificant, indicating no asymmetry effects in the direction of translation. In contrast to the matched trials, no effect of Trial or interaction involving Trial was found for grade 9 immersion pupils. Thus, the data are most consistent with a One-Store Different Encoding model.

Discussion. Both matched and mismatched latency data were most consistent with a One-Store Different Encoding model of bilingual semantic memory. Thus it appears that, within the range studied, neither duration nor intensity of exposure to the second language affected the structural organization of bilingual

Table 13

Median Latencies and Accuracy Scores for Grade 9 Immersion Students on the First and Second Half of the Trials of the Category Matching Task

Trial	Language Condition			
	EE	FF	EF	FE
	Match trials			
First Half				
Latency[a]	1.27	1.52	1.35	1.37
Accuracy[b]	92	91	95	93
Second Half				
Latency	1.20	1.25	1.32	1.40
Accuracy	97	95	99	89
	Mismatch trials			
First Half				
Latency	1.33	1.48	1.38	1.49
Accuracy	97	89	95	92
Second Half				
Latency	1.28	1.32	1.41	1.44
Accuracy	97	96	97	97

Note. EE = English-English; FF = French-French; EF = English-French; FE = French-English.

[a]Latency values reported in seconds. [b]Accuracy scores reported in percentages.

bilingual memory. Even among pupils in the early stages of second language acquisition (grade 6 extended French pupils), one memory store appeared to be utilized. One exception was found in the matched latency data of grade 9 immersion pupils. On the second half of the trials, the data were consistent with a Two-store Equal Encoding model. This pattern of results may reflect the use of a task-specific strategy by grade 9 immersion pupils, such that a temporary French store is utilized. When faced with a French and English word that matched, the student merely had to translate to or from the temporary French store and decide whether the words were from the same category. The fact that no effect for Half was found for grade 9 immersion students on the mismatched trials does not negate this possibility. Rather, when grade 9 immersion students realized that the word and its translated equivalent were not from the same category, they may have checked the original semantic network.

IV. GENERAL DISCUSSION

One of the main purposes in jointly examining the development of production, comprehension, and organization of memory among second language learners was to develop greater understanding of processing and representation. To summarize, category knowledge for both languages was represented in one memory store. Comparison of the relative rate of development of second language comprehension and production skills revealed that fewer years of exposure were required to achieve a point of bilingual balance on picture-word matching than on either picture or word naming. This finding is consistent with previous research, which indicates that a point of bilingual balance is reached earlier on comprehension than production tasks (e.g., Albert & Obler, 1978; Kolers, 1966a). A differential rate of development was noted also between the two production tasks. A point of balance was achieved following fewer years of exposure on the word-naming task than on the picture-naming task.

The discussion begins with an examination of the effects of exposure on the development of comprehension and production skills. Then, processing models accounting for differential rates of development of these skills are reviewed, as well as the influence of one individual difference variable, Verbal IQ. To conclude, the relationship between encoding speed and representation in two languages is discussed.

A. Effects of Exposure

French latencies. Based on the assumption that the more an individual is exposed to a second language the faster words in that language can be processed (e.g., Ervin, 1961a; Lambert, 1955;), two questions were raised: (a) Is there a decrement of similar magnitude in French latencies across grade level for immersion and extended French pupils; and (b) do immersion students with more intense exposure to French respond faster in French than extended French pupils of the same age? Support was found for these predictions in some of the tasks.

Comparison of French latencies within each instructional group revealed that the means for both groups decreased across grade. The decrement was significant for immersion and extended French pupils on picture naming and picture-word matching, but only for extended French pupils on word naming. Nonsignificant findings are important as they indicate that, among immersion pupils, word naming latencies were so fast by grade 6 that the increased exposure between grade 6 and 9 only minimally decreased response latencies on this simple task (Dornic, 1979; Magiste, 1979; Shannon, 1982). This may be a floor effect that holds for the high frequency words used in this task. Similar conclusions may not hold for low frequency words, complex phrases, or sentences.

In general, the results indicated that intensity of exposure did effect French response latencies. Grade 6 immersion pupils were faster than grade 6 extended French pupils on all three tasks. At the grade 9 level, the difference in French

latencies between instructional groups was signficant only for picture naming, although the means were in the predicted direction for the other tasks. The nonsignificant differences for word naming and picture-word matching most likely reflect floor effects rather than the fact that extended French pupils had "caught up" to immersion students by grade 9.

English latencies. The expected patterns for first language latencies differed markedly from those for the second language. Based upon prior research (Favreau & Segalowitz, 1982, 1983; Magiste, 1979, 1982; Potter et al., 1984), two questions were generated: (a) Do immersion pupils have longer latencies than extended French pupils at the same grade level; and (b) do latencies increase as duration of exposure to French increases for immersion and extended French pupils?

Effects associated with intensity of exposure were explored by comparing immersion and extended French pupils at each grade level. On word naming, grade 9 immersion pupils who had reached a point of bilingual balance on that task took marginally longer than English-dominant extended French pupils. In picture naming, immersion students took longer than extended French pupils, but the difference was not significant at grade 6 or 9. These nonsignificant findings do not support Magiste's (1982) research, but neither do they strongly refute her claims because the means were in the predicted direction. Also, Magiste's (1982) conclusions were based on a comparison of balanced and nonbalanced bilinguals. In the present study, immersion pupils had not reached a point of balance on picture naming. Thus, the

comparison was between two groups of nonbalanced bilinguals, a factor which potentially could reduce the contrast between the latencies for the two groups. On the comprehension task, the difference in means between instructional groups was nonsignificant. Given that immersion pupils had reached a point of balance by grade 6, these results do not support Magiste's previous findings.

Effects of duration of exposure differed for immersion and extended French pupils. Among immersion students, a marginal increase in English word-naming latencies associated with increasing grade level was found, which is consistent with Magiste's (1979, 1982) previous findings. In contrast, a marginally significant decrease was found for extended French pupils. In picture naming, a small and nonsignificant decrease in English latencies across grade level was found for immersion students. For extended French pupils, however, there was a larger decrease, which approached significance. The results for immersion pupils may reflect a trade-off between decreases in English latencies associated with age (e.g., Goodman, et al., 1984) and increases expected for duration of exposure to French (e.g., Magiste, 1979). Therefore, the picture-naming data do not necessarily contradict Magiste's previous findings, but they do not clearly support them either. Rather, they indicate how effects associated with duration and age may vary as a function of intensity of exposure. The picture-word matching data do not support Magiste's predictions. Across grade, there was a decrement of approximately the same magnitude for immersion and

extended French pupils (Table 6).

To summarize, the finding of a significant relationship between intensity and duration of exposure for one of the production tasks, word naming, partially supports Magiste's (1979, 1982) previous research. Grade 9 immersion pupils took marginally longer to read English words aloud than extended French pupils in the same grade. Among extended French pupils, there was a significant decrement in word-naming latencies across grade, whereas among immersion pupils there was marginally significant increase. This pattern of results suggests that first language word-naming latencies increase as a function of duration of exposure to the second language, only if the exposure is intensive and pupils are approaching a point of bilingual balance. The data from the other production task, picture naming, neither supports nor refutes Magiste's claims, as the means were in the predicted direction but the difference between groups was nonsignificant. The results from the comprehension task did not support Magiste's previous findings as there was no significant difference between the groups at either grade level. Also, for both groups response speed decreased across grade, although the difference was only marginally significant for immersion pupils.

Reduced automaticity is one factor postulated to account for longer first language latencies (Dornic, 1979; Favreau & Segalowitz, 1982, 1982; Magiste, 1979, 1982). As a result of increased exposure to the second language, the amount of exposure to the first language is substantially reduced, which results in

slower first language response speeds.

Interference between the competing language systems is a second factor assumed to account for longer first language latencies (Dornic, 1979; Magiste, 1984). For example, when a bilingual has to name a picture, he must choose between two competing responses, which places demands on controlled processing, thereby increasing the response latency (Shiffrin et al., 1981).

The factors of reduced automaticity and interlingual interference can be used to account for the finding that grade 9 immersion pupils took marginally longer than extended French pupils to read words in English and for a similar trend in the picture-naming data, although nonsignificant. Immersion students received more intense exposure to French than extended French pupils, but they also had less opportunity to practice their English language skills. Thus, the slower English response speeds found for immersion pupils may have been due to reduced automaticity. Also, by grade 9 immersion pupils were approaching a point of bilingual balance on both tasks, whereas extended French pupils were English-dominant. Thus, immersion students may have experienced more competition or interference between their two language systems, a factor which would account for slightly longer English latencies.

On the picture-word matching task, no increase in first language latencies was found for duration or intensity of exposure. Nonsignificant duration effects for immersion pupils may have occurred because subjects at both grade levels reached a

point of bilingual balance and Magiste's (1979) data suggest that once balance is achieved, further increases in first language latencies associated with increased exposure are minimal. Therefore, increases attributable to reduced automaticity and interlingual interference would be minimized, also. However, students with more intense exposure to French (immersion) did not take longer than students with minimal exposure (extended French) when responding in English. This result appears to contradict the expected effects of reduced automaticity and interlingual interference (Magiste, 1979, 1982). Examination of differences between the production and comprehension tasks provides one possible explanation for these results. In picture naming, subjects have to choose which of two possible words, one French and one English, is correct. In word naming, the choice is between two phonological codes. In contrast, picture-word matching does not necessarily involve a choice between languages. Rather, subjects have to decide whether the picture and word match, but the language they have to work in is determined by the word. Therefore, additional processing time attributable to interference between the language system is minimized in the picture-word matching task.

B. Verbal Intelligence

In the present study, the average Verbal IQ scores of immersion pupils were more than one standard deviation above the mean. Similar, but less striking findings have been reported in earlier studies using Canadian immersion students (Trites, 1981;

Swain & Lapkin, 1981). Therefore, the present sample may be characteristic of Canadian anglophone immersion students. Among extended French pupils, the mean Verbal IQ score was also more than one standard deviation above the mean. To the author's knowledge, there is no previous research on the Verbal IQ scores of extended French pupils. Thus, it cannot be determined whether the present sample is characteristic of Canadian anglophone students in extended French programs. The generalizability of the present findings may be limited by higher Verbal IQ scores.

Although examination of the relationship between Verbal IQ and the development of comprehension and production skills in a second language was exploratory, some tenative conclusions can be drawn with respect to the issue raised in question 5. In previous research, a relationship has been reported between Verbal IQ and the acquisition of literacy skills in the second language, but not between Verbal IQ and listening and speaking skills (Genesee, 1976; Genesee, Allister & Morin, 1974; Swain & Lapkin, 1981). In the present study, significant negative correlations were found between Verbal IQ and word naming and between Verbal IQ and picture-word matching, but not between Verbal IQ and picture naming. Thus, the results support previous research findings, although the tasks used assessed skills at a more elementary level.

In the bilingual literature, a distinction has been made between two aspects of linguistic skill: basic interpersonal communication skills and cognitive and academic aspects of language proficiency (Swain, 1981). Previous research indicates

that IQ is related to the acquisition of cognitive and academic uses of language proficiency, but not to basic interpersonal communicative skills. Interpersonal communication skills were not examined in the present study. However, the significant correlations between Verbal IQ and picture-word matching and word naming latencies supports previous findings, as both measure written/academic uses of language. To explore the relationship between IQ and the acquisition of second language skills, further research is required in which the range in IQ scores is much larger and subjects are grouped on the basis of their IQ scores.

C. Differential Rates of Skill Development

A point of bilingual balance was achieved by grade 6 and 9 immersion pupils on a test of comprehension, picture-word matching. On the production task assessing reading skills at the one-word level, a point of bilingual balance was achieved by grade 9 immersion students. On picture naming, a task assessing production of oral language, grade 9 immersion pupils had reached a point of balance. The French-English differential, however, was larger for picture than word naming even when scaling artifacts were taken into account. Two conclusions can be drawn from these results that relate directly to the issues raised in question six. First, for tasks involving single stimuli (one word, one picture, or one of each), comprehension may develop to a point of bilingual balance more rapidly than production. Second, a point of bilingual balance may develop more rapidly on tasks assessing the ability to read printed words aloud than on-

tasks measuring the ability to see a picture and produce the correct verbal label.

In the majority of prior studies, comprehension skills in a second language were found to develop prior to production skills (Albert & Obler, 1978; Kolers, 1963a; Ramirez & Politzer, 1978; Swain & Lapkin, 1981). Criticisms leveled at previous studies include the use of different stimulus items, with scant attention paid to equating items for length or frequency ratings, and different levels of complexity in the comprehension and production tasks (e.g. Dornic, 1979, Shannon, 1982). For these reasons the same single-word items were used in the comprehension and production tasks in the present study. The importance of these two methodological changes is clearly illustrated by comparing the present results to Magiste's (1979) finding that the development of word naming precedes comprehension. Given that Magiste used different stimulus items for all three tasks and did not assess comprehension skills at the one-word level, her results may be confounded and the pattern found in the present study may reflect more accurately the differential rate of development of production and comprehension at the one-word level.

It is important not only that the level of task difficulty (e.g., one-word level) be controlled, but also that careful analysis be made of task requirements (e.g., level of processing). This latter point is clearly illustrated by the data from the picture- and word-naming tasks, both measures of production. The results indicated that word naming in the second

language developed more rapidly than picture naming.

According to Snodgrass's (1984) model (Figure 1), word naming requires Level II processing, which includes recognition of the prototypical visual image and corresponding acoustic image. Picture naming requires processing to the semantic/propositional Level III. Thus, picture-naming latencies should be longer than word-naming latencies in both languages. Findings of previous research have been consonant with this model. Monolinguals have been found to name (read) words aloud some 0.2 s to 0.3 s faster than pictures of the same items (e.g., Huttenlocher & Kubicek, 1983; Potter & Faulconer, 1975; Snodgrass, 1984). A similar pattern of differentially longer picture- than word-naming latencies has been reported in the bilingual literature. For example, an average difference of 0.3 s was reported for a sample of Chinese-English adult balanced bilinguals (Potter et al., 1984), 0.28 s for adult balanced Hebrew-English bilinguals (Shannon, 1982) and 0.4 s for adolescent balanced German-Swedish bilinguals (Magiste, 1979). In the present study, picture-naming latencies were longer than word-naming latencies in both languages. In English, the difference was approximately 0.3 s. However in French, the picture-word differential was considerably larger for English-dominant extended French pupils.

The slower rate of development of second language picture-naming skills is accounted for by the fact that it is a case of varied mapping practice, whereas word naming exemplifies direct mapping (Shiffrin & Schneider, 1977). For example, when a

bilingual sees a picture of a dog, sometimes he will say "dog" and the other times "chien". However, everytime he sees the word "chien", he will say it. Thus, more exposure to the second language would be required to be able to name pictures as fast as read words.

D. Facilitation Effects

The three tasks assessing comprehension and production were presented once in French and English. The question of interest was whether cross-language facilitation effects would be found and whether there would be a larger facilitation effect when items were presented first in English. Among balanced bilinguals cross-language facilitation effects have been reported to be of the same magnitude regardless of which language the word appeared in first (Kolers, 1966b; Kolers & Gonzalez, 1980). In contrast, among nonbalanced bilinguals cross-language facilitation effects have been reported to be larger from the dominant to nondominant language (Ehri & Ryan, 1980). The finding of larger cross-language facilitation from the dominant to nondominant language is a pattern expected for nonbalanced bilinguals. The fact that a differential cross-language facilitation effect was found for grade 9 immersion pupils, who demonstrated equivalent French-English word-naming latencies, and for grade 6 and 9 immersion pupils, who demonstrated equivalent French and English picture-word matching latencies, is surprising. However, although the French-English differential was nonsignificant for these groups, response speed was still faster in English than

French.

Cross-task facilitation effects were found only for word naming. Receiving the comprehension task prior to word naming did not facilitate performance. Investigation of this issue was exploratory in nature. The results suggest that receiving the picture and word first did not facilitate performance on word naming.

It has been previously found that practice effects are larger in the nondominant language (Kolers, 1966b; Macnamara, 1967). In the picture-word matching task, words were repeated twice. Practice effects were larger for extended French than immersion students. Although the expected difference between facilitation effects for French and English did not reach significance, the means were in the predicted direction (Table 6). A greater reduction in French latencies across trials was found for extended French than immersion pupils at grade 6 and 9.

E. Organization of Bilingual Semantic Memory

In studies where category matching (Potter et al., 1984) and semantic verification tasks (Caramazza & Brones, 1980) have been utilized, consistent support for the one-store model of bilingual semantic memory has been reported. A category matching task was used in the present study, but two aspects of the design differed from previous studies. First, in the previous two studies, adult balanced bilinguals were examined, whereas in the present investigation subjects of different age levels with varying degrees of competency in the second language were

assessed. Second, in the present study, encoding speed in each language was assessed, whereas these data were not available in the previous studies. These two design modifications made it possible to assess not only the organizational structure of bilingual semantic memory, but also the relationship between encoding speed (processing) and memory organization (representation). Thus, the usual distinction between one- and two-store models was differentiated further in terms of four models of bilingual semantic memory: One-Store Different Encoding, One-Store Equal Encoding, Two-Store Different Encoding, and Two-Store Equal Encoding.

The data for all four groups were most consistent with the predictions of a One-Store Different Encoding model. Differential encoding speed was indicated by significantly longer French than English latencies. The finding of a nonsignificant difference between same and different language conditions suggested that a translation parameter was not operative. Given that no translation parameter was evident, the finding of a nonsignificant difference between the two different language condition (i.e., a comparison to test for possible asymmetry in the direction of translation) was expected.

The category matching task assessed processing at the semantic level. These data were consistent with a one-store model of bilingual semantic memory, which accords with Snodgrass' (1984) conclusion that semantic (Level III) tasks will support a one-store model of memory. Because a task eliciting processing at the lexical level (Level II) was not included, Snodgrass'

prediction that such tasks will support a two-store model cannot be addressed.

Inclusion of subjects varying with respect to duration and intensity of exposure to the second language, allowed for exploration of the structural change hypothesis (Gekoski, 1980; Potter et al., 1984). The hypothesis is that two separate stores are used initially, and then a gradual transition to one memory store occurs as competency in the second language approximates competency in the first. In accordance with the previous findings of Goodman et al. (1985) and Potter et al. (1984), no support was found for the structural change hypothesis. The present data indicate that even students with minimal exposure to the second language (grade 6 extended French pupils) utilized one memory store.

In the present study and in previous research (Goodman et al., 1985; Potter et al., 1984), only students who could read words in both languages were assessed. Modifying the category matching task so that words pairs are presented orally is one possible method for studying students at the earliest stages of second language acquisition and thereby adequately testing the predictions of the structural change hypothesis.

V. CONCLUSIONS

The results from the present study suggest that the original formulation of the bilingual semantic memory controversy is in need of revision. No longer should the focus be on determining whether bilinguals have one or two memory stores, but rather on specifying the conditions under which performance is language-dependent or language-independent (Glucksberg, 1984). The results of the present study demonstrate the utility of examining the developing relationship between production, comprehension, and memory organization. A point of balance was achieved earlier on the picture-word matching task than on either of the two production tasks. Among immersion pupils, English word-naming latencies increased as duration of exposure increased, but among extended French pupils latencies decreased. At grade 9, immersion pupils took longer than extended French pupils to name pictures, and read words aloud in English. Thus, the means for grade 9 pupils were in the direction predicted by Magiste (1979, 1982) on both production tasks, but not on the comprehension task. In the category matching task, which elicts semantic or language-independent processing, even at the early stages of second language acquisition, subjects stored semantic knowledge in one memory store. Thus, level of encoding skill did not appear to be related to the structual organization of semantic knowledge in memory.

VI. REFERENCES

Albert, M. L. & Obler, L. K. (1978). The bilingual brain. New York: Academic Press.

Anderson, J. R. (1976). Language, memory and thought. Hillsdale NJ: Erlbaum.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categoies: A replication and extension of the Conneticut category norms. Journal of Experimental Psychology Monographs, 80(3, Pt 4).

Ben-Zeev, S. (1977). Mechanisms by which childhood bilingualism affects understanding of language and cognitive structures. In P. A. Hornby, Bilingualism: psychological, social and educational implications. New York: Academic Press.

Bruck, M., Lambert, W. E., & Tucker, G. R. (1974). Bilingual schooling through the elementary grades: The St. Lambert project at grade seven. Language Learning, 24, 183-204.

Caramazza, A., & Brones, I. (1980). Semantic classification by bilinguals. Canadian Journal of Psychology, 34, 77-81.

Carey, S. T. (Ed.) (1974). Bilingualism, biculturalism and education. Edmonton: University of Alberta Press.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determinants of picture-naming latency. Quarterly Journal of Experimental Psychology, 25, 85-95.

Clark, H. H., & Clark, E. (1977). Psychology and language. New York: Hartcourt Brace Jovanovich.

Cummins, J. (1978). The cognitive development of children in immersion programs. Canadian Modern Language Review, 34,

855-883.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. Review of Educational Research, 49, 222-231.

Cummins, J., & Gulustan, M. (1975). Some effects of bilingualism, biculturalism, cognitive functioning. In S. T. Carey (Ed.) Bilingualism, biculturalism, and education. Edmonton: University of Alberta Press.

Dalrymple-Alford, E. C. (1968). Interlingual interference in a colour naming task. Psychonomic Science, 10, 215-216.

deVilliers, J. G., & deVilliers, P. A. (1978). Language acquisition. Cambridge, MA: Harvard University Press.

Dillon, R. F., McCormack, P. D., Petrusic, W. M., Cook, G. M., & Lafleur, I. (1973). Release from proactive interference in compound and coordinate bilinguals. The Bulletin of the Psychonomic Society, 4, 113-115.

Dornic, S. (1969). Verbal factors in number perception. Acta Psychology, 29, 393-396.

Dornic, S. (1977). Information processing and bilingualism. Unpublished manuscript. University of Stockholm.

Dornic, S. (1979). Information processing in bilinguals: Some selected issues. Psychological Research, 40, 329-348.

Dornic, S. (1982). Interlingual processing in verbal learning. Unpublished manuscript. University of Stockholm.

Dryer, F. N. (1971). Colour-naming interference in monolinguals and bilinguals. Journal of Verbal Learning and Verbal Behavior, 10, 297-302.

Dulay, H., Burt, M., & Krashen, S. (1982). Language two. New York: Oxford Press.

Durga, R. (1978). Bilingualism and interlingual interference. Journal of Cross Cultural Psychology, 9, 401-415.

Ehri, L. C. (1976). Do words really interfer in naming pictures. Child Development, 47, 502-505.

Ehri, L. C., & Ryan, E. B. (1980). Performance of bilinguals in a picture-word interference task. Journal of Psycholinguistic Research, 9, 285-302.

Ervin, S. M. (1961a). Semantic shift in bilingualism. American Journal of Psychology, 74, 233-241.

Ervin, S. M. (1961b). Learning and reading in bilinguals. American Journal of Psychology, 74, 446-451.

Faubert, J. (1979). Grammaire et orthographie 2e annee. Montreal: Guerin Ltd.

Favreau, M., Komoda, M. K., & Segalowitz, N. (1980). Second language reading: Implications for the word superiority effect in skilled bilinguals. Canadian Journal of Psychology, 4, 370-381.

Favreau, M., & Segalowitz, N. (1982). Second language reading in fluent bilinguals. Applied Psycholinguistics, 3, 329-34.

Favreau, M., & Segalowitz, N. (1983). Automatic and controlled processs in the first and second-language reading of fluent bilinguals. Memory and Cognition, 11, 565-574.

Gekoski, W. L. (1980). Language acquisition context and language organization in bilinguals. Journal of

Psycholinguistic Research, 9, 429-449.

Genesee, F. (1974). Evaluation of English-writing skills of students in French immersion programs. Report by Curriculum Department of the Protestant Board of Greater Montreal.

Genesee, F. (1976). The role of intelligence in second language learning. Language Learning, 26, 267-280.

Genesee, F., Allister, T., & Morin, S. (1974). Evaluation of the 1973-74 Pilot Grade XI French immersion class. Report by the Curriculum Department of the Protestant School Board of Montreal.

Genesee, F., & Hamayan, E. (1980). Individual differences in second language learning. Applied Psycholinguistics, 1, 95-110.

Genesee, F., & Lambert, W. E. (1983). Trilingual education for majority-language children. Child Development, 54, 105-114.

Genesee, F., Tucker, G. R., & Lambert, W. E. (1975). Communication skills of bilingual children. Child Development, 46, 1010-1014.

Glanzer, M., & Duarte, A. (1971). A repetition between and within languages in free recall. Journal of Verbal Learning and Verbal Behavior, 10, 625-630.

Glucksberg, S. (1984). Commentary: The functional equivalence of common and multiple codes. Journal of Verbal Learning and Verbal Behavior, 23, 100-104.

Goggin, J., & Wickens, D. D. (1971). Proactive interference and

language change in short-term memory. Journal of Verbal Learning and Verbal Behavior, 10, 453-458.

Goodman, G. S., Haith, M. M., Guttentag, R. E., & Rao, S. (1985). Automatic processing of word meaning: Intralingual and interlingual interference. Child Development, 56, 103-118.

Guenther, R. (1984). Bilingual representation of discourse. Psychological Reports, 54, 83-90.

Guttentag, R. E., Haith, M. M., Goodman, G. S., & Hauch, J. (1984). Semantic processing of unattended words by bilinguals: A test of the input switch. Journal of Verbal Learning and Verbal Behavior, 23, 178-188.

Harley, B., & Swain, M. (1977). An analysis of verb forms and function in the speech of French immersion pupils. Working Papers on Bilingualism, 14, 31-46.

Harley, B., & Swain, M. (1978). An analysis of the verb system used by young learners of French. Interlanguage Studies Bulletin, 3, 35-39.

Hatch, E. M. (Ed.) (1978). Second language acquistion: A book of readings. Rowley, MA: Newbury House.

Huttenlocher, J., & Kubicek, L. F. (1983). The source of relatedness effects on naming latency. Journal of Experimental Psychology, 9, 486-496.

Ianco-Worrall, A. D. (1972). Bilingualism and cognitive development. Child Development, 43, 1390-1400.

Immersion Bureau de l'education Francais (1980). Premiere annee. Manitoba: Ministere de l'education.

Kintsch, W. (1970). Recognition memory in bilingual subjects. Journal of Verbal Learning and Verbal Behavior, 9, 405-409.

Kintsch, W., & Kintsch, E. (1969). Interlingual interference and memory processes. Journal of Verbal Learning and Verbal Behavior, 8, 16-19.

Kiyak, M. A. (1982). Interlingual interference in naming colours. Journal of Cross-Cultural Psychology, 13,125-131.

Kolers, P. A. (1963a). Interlingual word associations. Journal of Verbal Learning and Verbal Behavior, 2, 291-300.

Kolers, P. A. (1963b). Bilingualism and bicodalism. Journal of Verbal Learning and Verbal Behavior, 2, 110-115.

Kolers, P. A. (1966a). Interlingual facilitation of short-term memory. Journal of Verbal Learning and Verbal Behavior, 5, 314-319.

Kolers, P. A. (1966b). Reading and talking bilingually. American Journal of Psychology, 79, 357-376.

Kolers, P. A. (1968). Bilingualism and information processing. Scientific American, 218, 78-86.

Kolers, P. A., & Gonzalez, E. (1980). Memory for words, synonyms and translations. Journal of Experimental Psyhology: Human Learning and Memory, 6, 53-65.

Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures and concepts: A comparison of lexical, object and reality decisions. Journal of Verbal Learning and Verbal Behavior, 23, 39-66.

Kucera, H., & Francis, W. N. (1967). Computation analysis of present day American English. Providence, RI: Brown

University Press.

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. Journal of Experimental Psychology, 99, 199-208.

Lambert, W. E. (1955). Measurement of linguistic dominance in bilinguals. Journal of Abnormal and Social Psychology, 50, 197-200.

Lambert, W. E., Havelka, J., & Gardner, R. D. (1959). Linguistic manifestations of bilingualism. American Journal of Psychology, 72, 77-82.

Lambert, W. E., Ignatow, M., & Krauthamer, I. (1968). Bilingual organization in free recall. Journal of Verbal Learning and Verbal Behavior, 7, 207-214.

Lambert, W. E., & Tucker, G. R. (1972). Bilingual education of children: The St. Lambert experiment. Rowley, MA: Newbury House.

Liepmann, D., & Saegert, J. (1974). Language tagging in bilingual free recall. Journal of Experimental Psychology, 103, 1137-1141.

Lopez, M., & Young, R. K. (1974). The linguistic independence of bilinguals. Journal of Experimental Psychology, 103, 981-983.

Lupker, S. J., & Katz, A. N. (1982). Can automatic picture processing influence word judgments? Journal of Experimental Psychology: Learning, Memory and Cognition, 8, 418-434.

Mackey, W. F. (1965). Bilingual interference: Its analysis and

measurement. Journal of Communications, 15, 239-249.

Macnamara, J. (1967). The effect of instruction in a weaker language. Journal of Social Issues, 23, 121-135.

Macnamara, J., Krauthamer, M., & Bolgar, M. (1968). Language switching in bilinguals as a function of stimulus and response uncertainty. Journal of Experimental Psychology, 78, 208-215.

Macnamara, J., & Kushnir, S. (1971). Linguistic independence of bilinguals: The input switch. Journal of Verbal Learning and Verbal Behavior, 10, 480-487.

McCormack, P. D. (1974). Bilingual linguistic memory: Independence or interdependence, two stores or one? In S. T. Carey (Ed.), Bilingualism, biculturalism and education. Edmonton: University of Alberta Press.

McCormack, P. D. (1976). Language as an attribute of memory. Canadian Journal of Psychology, 30, 238-248.

McCormack, P. D. (1977). Bilingual linguistic memory: The independence-interdependence issue revisited. In P. A. Hornby (Ed.), Bilingualism, psychological, social and educational implications. New York: Academic Press.

McCormack, P. D., & Colleta, S. P. (1975). Recognition memory for items from unilingual and bilingual lists. The Bulletin of the Psychonomic Society, 6, 149-157.

McLaughlin, B. (1977). Second language learning in children. Psychological Bulletin, 84, 438-459.

McLaughlin, B. (1978). Second language acquisition in childhood. Hillsdale, NJ: Erlbaum.

MacLeod, C. M. (1976). Bilingual episodic memory: Acquisition and forgetting. Journal of Verbal Learning and Verbal Behavior, 15, 347-364.

Magiste, E. (1976). Recall of concrete and abstract sentences in bilinguals. Scandinavian Journal of Psychology, 20, 179-185.

Magiste, E. (1979). The competing language systems of the multilingual: A developmental study of decoding and encoding processes. Journal of Verbal Learning and Verbal Behavior, 18, 79-89.

Magiste, E. (1982). Automaticity and interference in bilinguals. Psychological Research, 44, 29-43.

Magiste, E. (1984). Stroop tasks and dichotic translation: The development of interference patterns in bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 304-315.

O'Neill, W., & Huot, R. (1984). Release from proactive inhibition as a function of language of pronunciation shift in bilinguals. Canadian Journal of Psychology, 58, 54-62.

Oren, D. L. (1981). Cognitive advantages of bilingual children. Journal of Educational Research, 74, 163-169.

Pacella, R. G. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), Human information processing: Tutorials in performance and cognition. Hillsdale, NJ: Lawrence Erlbaum Associate Publishers.

Paivio, A., & Begg, I. (1981). Psychology of language.

Englewood Cliffs, NJ: Prentice Hall.

Paivio, A., & DesRochers, A. (1980). A dual approach to bilingual memory. Canadian Journal of Psychology, 34, 388-399.

Paivio, A., & Lambert, W. E. (1981). Dual coding and bilingual memory. Journal of Verbal Learning and Verbal Behavior, 20, 532-539.

Pellegrino, J.W., Rosinski, R.R., & Siegel, A.W. (1978). Picture-word differences in semantic decision latency: An analysis of single and dual-memory models. Memory and Cognition, 5, 81-98.

Potter, M. C., & Faulconer, B. A. (1975). Time to understand words and pictures. Nature, 253, 437-438.

Potter, M. C., So, K. F., VonEckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. Journal of Verbal Learning and Verbal Behavior, 23, 23-38.

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop colour-word task. Journal of Verbal Learning and Verbal Behavior, 8, 295-301.

Ratcliff, R. (1985). Theoretical interpretations of the speed-accuracy of positive and negative responses. Psychological Review, 92, 212-225.

Ramirez, A. G., & Politzer, R. L. (1978). Comprehension and production in English as a second language by elementary school children and adolescents. In E. M. Hatch (Ed.),

Second Language Acquisition. Rowley, MA: Newbury House.

Rosinski, R. R. (1977). Picture-word interference is semantically based. Child Development, 48, 643-647.

Rosinski, R. R., Golinkoff, R. M., & Kubkish, K. S. (1975). Automatic semantic processing in a picture-word interference task. Child Development, 46, 247-253.

Rosinski, R. R., Pellegrino, J. W., & Siegel, A. W. (1977). Developmental changes in the semantic processing of pictures and words. Journal of Experimental Child Psychology, 23, 282-291.

Saegert, J., & Young, R. K. (1975). Translation errors for abstract and concrete responses in a bilingual paired associate task. Bulletin of the Psychonomic Society, 6, 429.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. Journal of Verbal Learning and Verbal Behavior, 23, 84-99.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I. Detection, search and attention. Psychological Review, 84, 1-66.

Shannon, B. (1982). Identification and classification of words and drawings in two languages. Quarterly Journal of Experimental Psychology, 34, 135-152.

Shiffrin, R. M., Dumais, S., & Schneider, W. 9(1981). Characteristics of automatism. In L. Long & A. Baddeley (Eds.), Attention and performance. IX. Hillsdale,

NJ:Erlbaum.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing. II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84, 124-139.

Shoben, E. J. (1982). Semantic and lexical decisions. In C. R. Puff (Ed.), Handbook of research methods in human memory and cognition. New York: Academic press.

Smith, E. E. (1978). Theories of semantic memory. In W. K. Estes (Ed.), Handbook of learning and cognitive processes: Linguistic functions and cognitive theory. Hilldale, NJ: Lawrence Erlbaum Associates.

Snodgrass, J. G. (1984). Concepts and their surface representation. Journal of Verbal Learning and Verbal Behavior, 23, 3-22.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 20 pictures: Norms for familiarity and visual complexity. Journal of Experimental Psychology: Human Learning and Memory, 6, 174-215.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. Acta Psychologica, 30, 276-315.

Swain, M. (1981). Time and timing in bilingual education. Language Learning, 31, 1-16.

Swain, M., & Lapkin, S. (1981). Bilingual education in Ontario: A decade of research. Toronto: Ministry of Education.

Szamosi, M., Swain, M., & Lapkin, S. (1979). Do early immersion

pupils 'know' French? Orbit, 3, 20-23.

Taylor, I. (1971). How are words from two languages organized in a bilingual's memory? Canadian Journal of Psychology, 25, 228-239.

Taylor, I. (1976). Similarity between French and English words - A factor to be considered in bilingual language behavior. Journal of Psycholinguistic Research, 5, 25-94.

Trites, R. (1981). Primary French immersion: Disabilities and prediction of success. Review and Evaluation Bulletins, 2, Ontario Ministry of Education.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.) Organization and Memory, (pp. 381-403). New York: Academic Press.

Tulving, E., & Colota, A. (1970). Free recall of trilingual lists. Cognitive Psychology, 1, 86-98.

Weinreich, W. (1953). Languages in contact. New York: Lingustic Circle of New York.

Young, R. K., & Navar, M. I. (1968). Retroactive inhibition with bilinguals. Journal of Experimental Psychology, 77, 109-115.

Young, R. K., & Saegert, J. (1966). Transfer with bilinguals. Psychonomic Science, 6, 161-166.

Young, R. K. & Webber, A. (1967). Positive and negative transfer with bilinguals. Journal of Verbal Learning and Verbal Behavior, 6, 872-874.

VII. APPENDIX A: COMPREHENSION AND PRODUCTION STIMULI NORMS

For all words used in the picture-naming, word-naming, and picture-word matching tasks, the following normative data is listed below: two measures of name agreement the H index (Lachman, 1973) and percentage agreement (% Agree) (Snodgrass & Vanderwart, 1980); familiarity rating (FR) (Snodgrass & Vanderwart, 1980); Kŭcera-Francis (1967) frequency norms (K-F); age of acquisition (Age) norms (Carroll & White, 1973); and the French equivalent.

Table A1

Practice Words for the Picture-Naming, Word-Naming and Picture-Word Matching Tasks

English	H	% Agree	FR	K-F	Age	French
1. airplane	1.77	60	3.78	11	2.59	avion
2. barn	1.31	69	2.38	29	-	grange
3. bowl	.17	95	4.18	23	-	bol
4. cherry	.52	83	3.38	6	-	cherise
5. glasses	1.07	64	4.00	29	-	lunettes
6. gun	1.09	74	2.68	118	-	pistol
7. ladder	.16	98	3.35	19	3.12	échelle
8. mountain	.60	90	2.70	33	3.21	montagne
9. thumb	.16	98	4.72	10	-	pouce

Table A2

Test Words for Picture Naming, Word Naming, and Picture-Word Matching

English	H	% Agree	FR	K-F	Age	French
1. apple	.16	98	3.98	9	1.91	pomme
2. bed	.00	1.00	4.72	127	-	lit

Table A2 (continued)

English	H	% Agree	FR	K-F	Age	French
3. bell	.00	100	2.20	18	2.36	cloche
4. belt	.16	98	4.12	29	-	ceinture
5. bird	.69	88	3.62	31	-	oiseau
6. book	.00	100	4.75	193	1.83	livre
7. bottle	.28	95	3.72	76	-	bouteille
8. box	.80	88	2.88	70	-	boîte
9. cake	.84	83	4.02	13	2.06	gateau
10.car	1.08	81	4.70	274	-	voiture
11.chair	.00	100	4.58	66	1.86	chaise
12.church	.44	93	2.28	348	2.62	église
13.clock	.16	98	4.38	29	2.36	horloge
14.cloud	.17	95	3.82	28	-	nuage
15.coat	.95	79	3.88	43	-	manteau
16.cup	.44	93	4.49	45	1.66	tasse
17.desk	.32	95	4.32	65	-	pupitre
18.door	.16	98	4.68	312	1.97	porte
19.finger	1.37	71	4.78	49	-	doigt
20.fish	.00	100	3.28	35	2.61	poisson
21.flag	.32	95	2.90	16	-	drapeau
22.flower	.48	93	3.88	23	2.15	fleur
23.fork	.00	100	4.78	14	2.24	fourchette
24.glass	.16	98	4.78	99	-	verre
25.glove	.16	98	3.38	9	3.12	gant

Table A2 (continued)

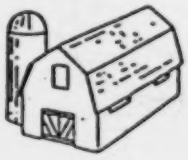
English	H	% Agree	FR	K-F	Age	French
26. hammer	.00	100	3.48	9	3.50	marteau
27. hat	.16	98	3.18	56	-	chapeau
28. heart	.00	100	3.72	173	-	coeur
29. house	.32	95	4.38	591	-	maison
30. key	.00	100	4.85	88	3.38	clé
31. knife	.00	100	4.85	88	3.38	couteau
32. leaf	.53	90	4.30	12	2.61	feuille
33. leg	1.11	81	4.65	58	-	jambe
34. lemon	.00	100	3.25	18	3.06	citron
35. moon	1.68	63	3.98	60	-	lune
36. peach	1.19	74	2.90	3	2.79	pêche
37. pen	.32	95	4.78	18	-	stylo
38. shirt	.00	100	4.56	27	-	chemise
39. shoe	.28	95	4.62	14	1.94	soulier
40. skirt	.16	98	3.64	21	-	jupe
41. star	.00	100	3.35	25	-	étoile
42. suitcase	1.01	79	3.65	20	-	valise
43. sun	.00	100	4.90	112	-	soleil
44. sweater	.98	83	4.48	14	-	chandail
45. tree	.00	100	4.68	59	2.03	arbre
46. truck	.53	90	4.02	57	-	camion
47. wheel	.33	93	2.22	56	-	roue
48. window	.32	95	4.40	119	2.28	fenêtre

VIII. APPENDIX B: PICTORIAL STIMULI

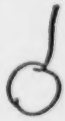
Pictures for the practice and test trials for picture naming and picture-word matching are shown below with their identifying number. They are arranged in alphabetical order according to each picture's most common English name. The names for each picture and the norms can be found in Appendix A.

Pictures for Practice Trials

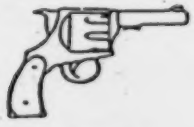
1.

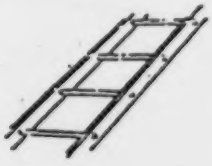
2. 

3.

4. 

5.

6. 

7. 

8.

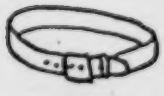
9.

Pictures for Test Trials

1.

2.

3.

4. 

Appendix B (continued)

5.

6.

7.

8.

9.

10.

11.

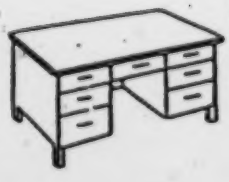
12.

13.

14.

15.

16.

17. 

18.

19.

20.

Appendix B (continued)

21.

22.

23.

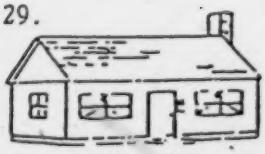
24.

25.

26.

27.

28.

29. 

30.

31.

32.

33.

34.

35.

36.

Appendix B (continued)

37.

38.

39.

40.

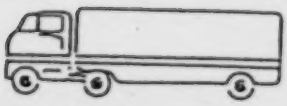
41.

42.

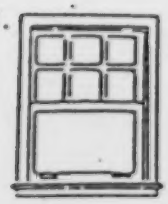
43.

44.

45.

46. 

47.

48. 

IX. APPENDIX C: WORD NORMS FOR CATEGORY MATCHING

The following information regarding the words used in the category-matching task is shown below: each English word and the name agreement, (H) (Lachman, 1973) and percentage agreement (% Agree) (Snodgrass & Vanderwart, 1980); familiarity rating (FR) (Snodgrass & Vanderwart, 1980); Kŭcera-Francis (1967) frequency norms (K-F); age of acquisition norms, (Age) (Carroll & White, 1973); and the French word equivalent.

Table C1

Practice Words

Category	English	H	% Agree	FR	K-F	Age	French
Clothing	belt	.16	98	4.12	29	-	ceinture
	coat	.95	79	3.88	43	-	manteau
	shirt	.00	100	4.56	27	-	chemise
	shoe	.28	95	4.62	14	1.94	solier
Fruit	apple	.16	98	3.98	9	1.91	pomme
	lemon	.00	100	3.25	18	3.06	citron
	pear	.00	100	3.55	6	-	poire
	peach	.19	74	2.90	3	2.79	pȩche

Table C2

Test Words

Category	English	H	% Agree	FR	K-F	Age	French
Parts of	arm	.53	90	4.74	94	-	bas
Human Body	ear	.28	95	4.50	29	1.82	oreille
	eye	.16	98	4.88	122	2.00	oeil

Table C2 (continued)

Category	English	H	% Agree	FR	K-F	Age	French
Human	foot	.28	95	4.78	70	-	pied
Body	hair	.64	90	4.59	148	-	cheveux
Parts	nose	.16	98	4.52	148	-	nez
	lips	.44	93	4.50	69	-	bouche
Animals	chicken	1.35	67	2.42	37	-	poulet
	cow	.44	93	2.43	29	1.90	vache
	dog	.00	100	4.60	75	1.55	chien
	horse	.00	100	3.55	117	2.67	cheval
	rabbit	.00	100	2.95	11	2.61	lapin
	pig	.60	90	2.18	8	2.94	cochon
	sheep	.95	67	1.85	23	-	mouton

X. APPENDIX D: WORD LISTS FOR CATEGORY MATCHING

Listed below are the four stimulus sets for the True trials and the four stimulus sets for the False trials. There are an equal number of words from each of two categories, animals and body parts. Eight words appear a total of 11 times in the 80 trials (i.e., arm, foot, lhand, lips, dog, pig, cow, and rabbit). Six words appear 12 times (i.e., ear, eye, nose, chicken, horse, and sheep). Word position was controlled, also.

Table D1

Word Pairs for the Category Matching Task

Stimulus Set	Item	Word Pairs	
1 (True)	1	arm	ear
	2	ear	eye
	3	eye	foot
	4	foot	hair
	5	hair	nose
	6	sheep	chicken
	7	chicken	dog
	8	dog	cow
	9	cow	horse
	10	horse	rabbit
2 (True)	11	arm	eye
	12	ear	foot
	13	eye	hair
	14	nose	foot
	15	hair	lips
	16	sheep	dog
	17	cow	chicken
	18	dog	horse
	19	cow	rabbit
	20	horse	pig
3 (True)	21	arm	foot
	22	hair	ear
	23	eye	nose
	24	lips	arm
	25	nose	lips
	26	sheep	cow
	27	chicken	horse
	28	rabbit	dog
	29	pig	sheep
	30	rabbit	pig
4 (True)	31	arm	hair
	32	nose	arm
	33	ear	nose

Appendix D1 (continued)

Stimulus Set	Item	Word Pair	
	34	eye	lips
	35	lips	ear
	36	horse	sheep
	37	rabbit	sheep
	38	chicken	rabbit
	40	pig	chicken
5 (False)	41	foot	sheep
	42	chicken	ear
	43	eye	pig
	44	cow	arm
	45	hair	horse
	46	rabbit	nose
	47	lips	dog
	48	chicken	foot
	49	eye	cow
	50	pig	ear
6 (False)	51	pig	foot
	52	ear	sheep
	53	chicken	eye
	54	foot	rabbit
	55	arm	horse
	56	cow	hair
	57	hair	rabbit
	58	horse	lips
	59	arm	sheep
	60	dog	nose
7 (False)	61	foot	cow
	62	horse	ear
	63	rabbit	eye
	64	arm	chicken
	65	dog	hair
	66	nose	pig
	67	sheep	lips
	68	ear	dog
	69	sheep	nose
	70	lips	chicken
8 (False)	71	horse	foot
	72	ear	rabbit
	73	dog	eye
	74	lips	pig
	75	sheep	hair
	76	nose	chicken
	77	cow	lips
	78	eye	horse
	79	pig	arm
	80	nose	cow

XI. APPENDIX E: ACCURACY DATA

Picture Naming

The percentage of picture named correctly in French and English were the dependent measures. In English, a floor effect was obtained for all groups. In French, the range was from 63%, grade 6 extended French pupils, to 96%, grade 9 immersion students. Therefore, only French accuracy scores were analyzed.

One correlation with Verbal IQ was significant, $r(24) = .45$, $p < .05$. Thus, a Grade (six, nine) by Program (immersion, extended French) by sex (male, female) by Task Order (PI, WI, PWI) by Testing Sequence (English-French, French-English) analysis of covariance was computed with Verbal IQ as the covariate. The assumption of homogeneity of regression coefficients was met. The dependent measure was distributed binominally, and so an arcsin transformation was applied, prior to data analysis.

Main effects were found for Program, $F(1,45) = 15.17$, $p < .001$, Grade, $F(1,45) = 13.45$, $p < .001$, and Task Order, $F(2,45) = 25.88$, $p < .001$. Program, Grade, and Task Order interacted, $F(2,45) = 3.37$, $p < .05$. Examination of the simple main effects, revealed that for all four groups, subjects named more pictures correctly when they received picture-word matching prior to picture naming, $ps < .05$. All subjects, except grade 9 immersion students, $F(1,45) < 1$, named more pictures correctly when word naming preceded picture naming $ps < .05$.

These results indicate that accuracy was enchanced when the

the picture-word matching task preceded the picture-naming task. However, for all but the most advanced group of second language learners, grade 9 immersion studetns, merely seeing the word in the word-naming task prior to naming the picture, improved accuracy.

Word Naming

Percentages of words named correctly in French and English were the dependent measures. A floor effect was obtained in English. Variation in French was minimal: 100% for both immersion groups, and 86% and 96% for grade 6 and 9 extended French pupils, respectively. No further analyses were computed.

Picture-Word Matching

Examination of the percentage of picture-word pairs sorted correctly in English revealed nearly perfect scores for all groups. A floor effect was evident also in French for immersion students. Among extended French pupils, accuracy scores ranged from 84% to 95% for grade 6 and 9, respectively. No further analysis of the data was undertaken.

Category Matching

The range in accuracy scores was from 81% to 94%. None of the correlations between accuracy scores and Verbal IQ were significant. Thus, a Grade (six, nine) by Program (immersion, extended French) by Sex (male, female) by Half (first half, second half) by Language Condition (English-English, French-French, English-French, French-English) analysis of variance was computed with repeated measures on the last two factors. In line with the procedures followed for the analysis of latency data, match and mismatch accuracy scores were analyzed

separately.

For the matched data, main effects were found for Program, $F(1,92) = 25.53$, $p < .001$, Half, $F(1,92) = 9.3$, $p < .01$, and Language Condition, $F(3, 276) = 13.22$, $p < .001$. French immersion pupils were more accurate than extended French pupils, and accuracy scores were higher on the second than first half of the trials.

Language Condition interaction with Half, $F(3,276) = 5.48$, $p < .01$. Tests of simple main effects revealed that, although there was an increase in accuracy across Half for English-English, French-French, and English-French word pairs, only the increase for English-French word pairs was significant, $p < .05$.

Language Condition interacted with Program, $F(3,276) = 3.04$, $p < .03$. Immersion pupils had higher accuracy levels than extended French pupils at each language condition, but the differences were significant only for French-English and English-French word pairs, $ps < .05$.

Main effects identified on matched trials were replicated for mismatched trials: Program, $F(1,92) = 16.75$, $p < .001$, Half, $F(1,92) = 6.92$, $p < .01$, and Language Condition, $F(3,276) = 10.21$, $p < .001$.

Program and Language Condition interacted, $F(3.276) = 4.02$, $p < .03$. Examination of the simple main effects indicated that immersion students made fewer errors than extended French pupils on French-French, English-French, and French-English word pairs, $ps < .01$. The difference for English-English word pairs was nonsignificant, $F(1,276) < 1$. Students in both programs made more errors on

French-French than English-English word pairs, $ps < .01$. However, no significant difference was found between the two different language conditions, English-French and French-English, for immersion, $F(1,276) < 1$, or extended French pupils, $F(1,276) < 1$. The difference between same and different language conditions was not significant for either immersion, $F(1,276) < 1$, or extended French pupils, $F(1,276) < 1$.